

Does Bigger Data Make Better Plans? Decomposing the Value of Big Data Analytics in Strategic and Systematic Production Planning under Demand Uncertainty

Eleanor Hawking^{a,*}

^a Institute for Environmental Science, University of Bristol, UK

* e.hawking@bristol.ac.uk

ABSTRACT

Big data analytics (BDA) is widely argued to improve strategic and systematic planning, yet the empirical evidence relates analytics capability to firm performance without isolating which component of an analytics system actually creates the value. This study decomposes that value for capacity and production planning under demand uncertainty. Seven planning regimes, ranging from judgmental annual planning to a full BDA regime combining cross-series machine learning, an external leading indicator, probabilistic forecasts and monthly revision, are embedded in an identical rolling-horizon two-stage stochastic planning model and evaluated on 200 real monthly industrial demand series (180 months each) drawn from the M4 competition. Because execution is identical across regimes, differences in realised performance are attributable solely to planning information. The BDA regime reduced planning-relevant cost per unit by 30.4% relative to judgmental planning and by 11.9% relative to quarterly statistical planning, and captured 70.7% of the gap between judgmental planning and a perfect-information bound. Decomposition shows that formalising the forecast contributes 48.8% of that gap, big-data forecast quality a further 21.0%, and representing uncertainty explicitly 13.8%, whereas increasing the revision frequency from quarterly to monthly does not help and slightly harms performance while increasing plan instability. Superior forecasts used deterministically were statistically indistinguishable from a simple quarterly statistical plan. The advantage of BDA increased with demand volatility (4.8% to 17.2% across volatility terciles) and with the cost of shortages, and decreased when capacity lead times exceeded the horizon over which the external signal was informative. The results indicate that analytics investment should prioritise decision-quality elements, namely uncertainty representation and cost-aware optimisation, over data velocity.

Paper type: Research

Received 2026-04-15

Revised 2026-05-30

Accepted 2026-06-11

Keywords:

Big data analytics;

Strategic planning;

*Production and capacity
planning;*

Demand uncertainty;

Stochastic programming;

Simulation

1. Introduction

Strategic and systematic planning determines how an industrial firm commits capacity, capital and material ahead of uncertain demand. These commitments are costly to reverse, and their quality is bounded by the information available when they are made. The Industry 4.0 transition has multiplied that information: firms now hold high-frequency transactional records, sensor streams and external indicators that were unavailable when classical planning methods were developed [1], [2], and the emerging Industry 5.0 agenda adds resilience and sustainability to the objectives such plans must serve

[3]. Big data analytics (BDA) is accordingly promoted as a means of making planning more responsive, more granular and more evidence-based [4].

The empirical literature broadly supports a positive association between analytics capability and firm or supply-chain performance [5], [6], [7], [8]. That literature, however, is predominantly survey-based and treats analytics as a composite capability. It therefore cannot answer the question a planning manager actually faces: given a limited budget, which component of an analytics system should be improved first? An analytics upgrade bundles at least three distinct changes. It broadens the data used, from a plant's own aggregate history to many related series and external signals. It changes the representation of uncertainty, from a single point forecast to a full predictive distribution. And it increases the revision frequency of the plan, from annual or quarterly cycles towards near-continuous replanning. These components have different costs, different organisational implications, and, as this paper shows, very different returns.

Separating them with field data is difficult, because firms rarely change one component at a time and the counterfactual plan is never observed. This study therefore uses a controlled simulation in which the components can be varied independently while everything else is held fixed. To avoid the artificiality of a synthetic demand process, the simulation is driven by 200 real monthly industrial demand series from the M4 forecasting competition [9]. Seven planning regimes, spanning judgmental, statistical and big-data planning, are embedded in an identical rolling-horizon two-stage stochastic capacity-planning model [10], and all regimes are executed under an identical operating policy. Any difference in realised cost, service or plan stability is thus attributable only to the planning information.

Three research questions are addressed. RQ1: how much does a big-data planning regime improve realised planning performance relative to judgmental and conventional statistical planning, and how much of the theoretically achievable improvement does it capture? RQ2: how is that improvement distributed across the three components of the analytics upgrade, namely data breadth, uncertainty representation and revision frequency? RQ3: under which conditions, in terms of demand volatility, capacity lead time and shortage cost, is BDA most and least valuable?

The study contributes in four ways. First, it provides an experimental decomposition of BDA value in planning, rather than an aggregate capability-performance association, and identifies which component dominates. Second, it reports a result that qualifies the prevailing emphasis on data velocity: increasing revision frequency from quarterly to monthly did not improve performance and increased plan instability, whereas representing uncertainty explicitly was consistently valuable. Third, it shows that forecast accuracy alone is not sufficient: a more accurate forecast used deterministically performed no better than a far simpler quarterly statistical plan. Fourth, it maps the contingencies that govern the return on analytics investment, which gives managers a basis for deciding where analytics will pay. Section 2 reviews the literature, Section 3 describes the method, Section 4 reports results, Section 5 discusses implications and limitations, and Section 6 concludes.

2. Literature Review

2.1. Big data analytics and organisational performance

A substantial body of research links analytics to performance. Wamba et al. [5] found that BDA capability affects firm performance through process-oriented dynamic capabilities, and Gunasekaran et al. [6] reported similar effects for supply-chain and organisational performance. Chae et al. [7] adopted a resource-based view and showed that supply-chain analytics is associated with better operational

performance, while Mikalef et al. [8] found that the effect on innovation is mediated by dynamic capabilities and moderated by environmental conditions. Reviews of analytics in operations management emphasise the breadth of potential applications but also note the scarcity of studies that quantify operational consequences rather than perceptions [4]. At the macro level, the gap between the expected and measured returns to data-driven technologies remains an open issue [11], [12].

Two features of this literature limit its guidance for planning practice. It generally measures analytics as a latent construct assembled from perceptual items, so the specific mechanism that creates value is not identified; and performance is usually self-reported rather than computed from realised operational outcomes. As a result, the literature establishes that analytics matters without establishing which part of it matters.

2.2. Analytics in production and capacity planning

Planning research has long recognised that decisions under uncertainty should be based on distributions rather than point estimates [13], [14], and stochastic programming provides the formal machinery for such decisions [10], [15]. In parallel, the predictive-to-prescriptive stream shows how rich covariate data can be mapped directly into decisions, with performance measured by decision cost rather than prediction error [16]. Forecasting research has demonstrated that external macroeconomic indicators can improve tactical sales forecasts when suitable regularisation is used [17], and large-scale competitions have shown that methods learning across many series can outperform purely local models [9]. Hybrid approaches combining learning with optimisation have been applied to demand prediction, portfolio and scheduling problems in industrial settings [18], [19], [20], [21], [22], and simulation-based optimisation has been used where analytical evaluation of a policy is intractable [23], [24].

These strands are rarely combined. Forecasting studies typically stop at accuracy metrics; planning studies typically assume a given demand distribution; and analytics-performance studies measure neither. Consequently the operational value of better planning information, and its distribution across components, remains largely unquantified.

2.3. Judgmental planning as the practical baseline

In many firms, strategic capacity plans are still set judgmentally in annual cycles. Behavioural research documents systematic distortions in such forecasts, including anchoring on recent outcomes and insufficient adjustment [25], and reviews of judgmental forecasting report persistent optimism bias and inconsistency, alongside conditions under which judgment adds value [26]. A meaningful evaluation of analytics should therefore benchmark against this judgmental baseline rather than only against a statistical one, since the practical decision facing many firms is whether to move away from judgment at all.

Table 1. Positioning of this study relative to representative literature.

Stream	Typical design	Performance measure	Component of analytics isolated	Examples
Analytics capability and firm performance	Cross-sectional survey, SEM	Perceived firm/SC performance	None (composite construct)	[5], [6], [7], [8]
Forecasting methods	Benchmark competition on many series	Forecast error	Forecast model only	[9], [17], [27]

Stream	Typical design	Performance measure	Component of analytics isolated	Examples
Stochastic and prescriptive planning	Optimisation with assumed distributions	Expected decision cost	Decision model only	[10], [16], [15]
Hybrid learning and optimisation in industry	Case study, algorithm design	Objective value of the model	Solution method	[18], [21], [22]
This study	Controlled simulation on 200 real demand series	Realised planning cost, service, stability	Data breadth, uncertainty, revision frequency, separately	—

2.4. Research gap

Three gaps follow. First, the value of BDA in planning has not been decomposed into the components that an investment decision must choose between. Second, planning studies seldom benchmark against both judgmental practice and a perfect-information bound, so the share of the achievable improvement that analytics actually captures is unknown. Third, the contingencies that govern the return on analytics, in particular demand volatility and the structural parameters of the capacity decision, have not been mapped in an operational setting. This study addresses these gaps.

3. Methodology

3.1. Framework and experimental logic

Figure 1 presents the planning cycle used throughout. A firm senses demand information, produces a forecast, decides capacity and production commitments under an acquisition lead time, and executes. Realised performance feeds back through the revision cycle. An analytics upgrade is modelled as movement along three independent dimensions: data breadth (own history versus many series plus external signals), uncertainty representation (point forecast versus full predictive distribution), and revision frequency (annual, quarterly or monthly). The experimental logic is to vary these dimensions while holding the demand data, the planning optimisation model, the cost structure and the execution policy constant, so that realised differences are caused only by planning information.

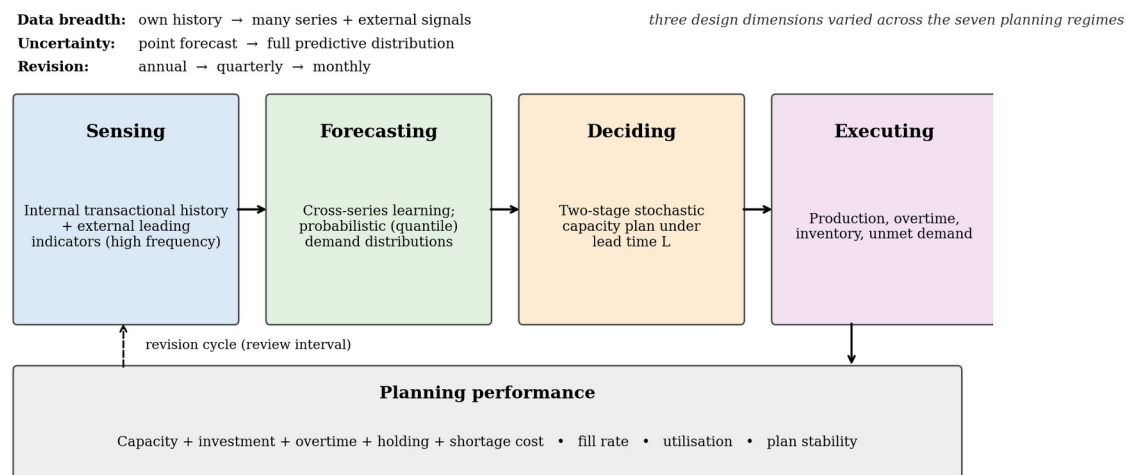


Figure 1. Planning cycle and the three design dimensions along which the planning regimes differ.

3.2. Demand data

Demand paths are real rather than simulated. From the M4 competition dataset [9], all monthly series classified as Industry were extracted, and those with at least 180 observations and strictly positive values were retained (5,669 series). A random sample of 200 series (seed fixed for reproducibility) was drawn, and the last 180 monthly observations of each were used. The first 120 months constitute history available before the experiment; the following 60 months are the simulated planning window. Table 2 summarises the panel. The series are heterogeneous in scale, trend and volatility, which is essential for testing whether cross-series learning transfers.

Table 2. Characteristics of the demand panel (200 monthly industrial series, 180 months each).

Characteristic	Mean	Median	10th pct.	90th pct.
Mean demand level (units/month)	—	3,849	1,304	9,472
Coefficient of variation	0.196	0.165	0.065	0.376
Trend strength (STL)	0.85	0.92	—	—
Seasonal strength (STL)	0.29	0.15	—	—
Growth, months 61–180 vs. 1–60 (%)	—	17.2	–17.4	73.1

3.3. Planning regimes

Seven regimes are compared (Table 3). The judgmental regime (JB) revises annually and forms forecasts by anchoring on the recent level, extrapolating a damped trend and applying an optimism bias, reproducing documented behavioural patterns [25], [26]. The statistical regime (TS) revises quarterly using a per-series damped-trend exponential smoothing model with additive seasonality [28]. TS-S is identical but passes a predictive distribution rather than a point forecast to the planning model. The big-data regimes use a single global gradient-boosted quantile model [29], [30] trained across all series, with scale-free features and an external leading indicator. BDA-Q revises quarterly with a distribution; BDA-D revises monthly with a point forecast; BDA revises monthly with a distribution. A perfect-information regime (PI) that plans on realised demand provides an upper bound on achievable performance.

Table 3. The seven planning regimes as a design matrix.

Regime	Forecast basis	Data breadth	Uncertainty	Revision
JB — judgmental	Anchor + damped trend + optimism bias	Own history	Point	Annual
TS — statistical	Damped-trend ETS with seasonality	Own history	Point	Quarterly
TS-S — statistical, stochastic	Damped-trend ETS with seasonality	Own history	Distribution	Quarterly
BDA-Q — big data, quarterly	Global gradient-boosted quantiles	All series + external signal	Distribution	Quarterly
BDA-D — big data, deterministic	Global gradient-boosted quantiles	All series + external signal	Point	Monthly
BDA — full big data	Global gradient-boosted quantiles	All series + external signal	Distribution	Monthly
PI — perfect information	Realised demand (bound)	—	—	Monthly

3.4. Forecast generation

All regimes forecast a horizon of $H = 18$ months. Forecasts made with information through month t cover months $t+1$ to $t+H$, and no feature uses information after t . The global model predicts demand scaled by the local 12-month mean at the review month, which makes one model transferable across series of different magnitudes; features comprise recent lags, moving averages, a short-term trend, a same-month seasonal reference from one or two years earlier, a volatility measure, the horizon and calendar month, and the external signal. Separate models are estimated for the 0.1, 0.3, 0.5, 0.7 and 0.9 quantiles [30] and refitted periodically using only data available at that time. The external leading indicator is a synthetic signal observable at month t whose correlation ρ with standardised demand over the following six months is a controlled parameter ($\rho = 0.5$ in the base case), representing external data of realistic but imperfect quality [17]; $\rho = 0$ is examined as an ablation. Predictive intervals for TS and TS-S are calibrated empirically from that series' own past relative forecast errors at origins preceding the simulation window, so that neither regime is advantaged by better-calibrated uncertainty. Forecast accuracy is measured by the mean absolute scaled error (MASE) relative to the in-sample seasonal-naive benchmark [27], and interval quality by empirical coverage [31].

3.5. Planning model

At each review, the plan solves a two-stage stochastic linear program by sample average approximation [10], [15]. First-stage variables are capacity additions a_k and retirements r_k , which take effect after a lead time L and are common to all scenarios. Second-stage variables are production p , overtime o , inventory I and sales per scenario. With S demand scenarios of probability π_s taken from the forecast quantiles, the program minimises

$$\min \sum_k [c_E a_k + c_R r_k + c_F K_k] + \sum_s \pi_s \sum_k [c_p p_{\{sk\}} + c_o o_{\{sk\}} + h I_{\{sk\}} - b \cdot \text{sales}_{\{sk\}}] \quad (1)$$

subject to the capacity and inventory balances

$$K_k = K_{\{k-1\}} + a_{\{k-L\}} - r_{\{k-L\}} + \text{pipeline}_k \quad (2)$$

$$I_{\{sk\}} = I_{\{s,k-1\}} + p_{\{sk\}} + o_{\{sk\}} - \text{sales}_{\{sk\}} \quad (3)$$

$$p_{\{sk\}} \leq K_k, \quad o_{\{sk\}} \leq \alpha K_k, \quad \text{sales}_{\{sk\}} \leq d_{\{sk\}}, \quad \text{all variables} \geq 0 \quad (4)$$

Only the first-period decisions are implemented, and the problem is re-solved at the next review, giving a rolling-horizon policy. Rewarding sales at rate b is equivalent to penalising unmet demand, so b is the shortage penalty. Execution is then identical for every regime: the firm produces towards the expected demand of the current month plus the planned inventory target, subject to available capacity and an overtime allowance, and unmet demand is lost. Base-case parameters are given in Table 4; costs are normalised so that unit production cost equals one.

Table 4. Base-case model parameters (costs normalised to unit production cost).

Parameter	Symbol	Value	Interpretation
Planning horizon	H	18 months	Rolling horizon of each plan
Capacity lead time	L	6 months	Delay before capacity changes take effect
Unit production cost	c_p	1.0	Numéraire
Overtime cost	c_o	1.5	Premium for capacity-exceeding output

Parameter	Symbol	Value	Interpretation
Overtime allowance	α	0.20	Overtime capped at 20% of capacity
Capacity carrying cost	c_F	0.15 per unit per month	Cost of holding installed capacity
Capacity expansion / retirement	c_E / c_R	3.0 / 1.0	Adjustment costs
Inventory holding cost	h	0.02 per unit per month	Carrying finished goods
Shortage penalty	b	3.0	Contribution and goodwill lost per unit

3.6. Performance measures and experimental design

The primary measure is planning cost per unit of demand: the sum of capacity carrying, capacity adjustment, overtime premium, inventory holding and shortage costs, divided by total demand. Standard production cost is excluded because the units produced must be produced under any regime, so including it would dilute differences that planning can actually influence; total cost including production is also reported. Secondary measures are the fill rate, capacity utilisation, cumulative capacity churn (total capacity adjusted, relative to average capacity) and plan nervousness, defined as the mean absolute change between successive capacity adjustments and normalised by mean monthly demand.

Each of the 200 series is simulated under all seven regimes over 60 months, giving 1,400 paired observations. Because regimes are compared on identical series, comparisons are paired: the Friedman test assesses overall differences and Wilcoxon signed-rank tests assess pairwise differences [32]. Sensitivity analyses vary the capacity lead time ($L = 3, 6, 12$ months) and the shortage penalty ($b = 1.5, 3.0, 6.0$) on the first 100 series, and results are additionally stratified by demand-volatility tercile. An ablation with $p = 0$ isolates the contribution of the external signal. All experiments use fixed random seeds and the implementation uses scikit-learn [33].

4. Results

4.1. Forecast accuracy

Table 5 and Figure 2 report forecast performance. Judgmental forecasts were far less accurate than either model-based regime (MASE 2.866), and were worse than a seasonal-naïve benchmark, consistent with documented biases in unaided judgment [26]. The global big-data model was more accurate than the per-series statistical model (MASE 0.812 versus 0.932, a 12.9% reduction), and its advantage was concentrated at short horizons (0.595 versus 0.701 for the first six months), which is where the external leading indicator is informative. Figure 2(b) shows this pattern across the horizon. Empirical coverage of the 80% predictive interval was close to nominal for both BDA and TS, confirming that the subsequent planning comparison is not driven by one regime having better-calibrated uncertainty. The judgmental regime, which supplies no distribution, is severely miscalibrated by construction.

Table 5. Forecast accuracy and interval calibration (200 series, 60 monthly origins, horizon 18).

Regime	MASE (all horizons)	MASE (horizons 1–6)	Below 10th pct. (target 0.10)	Above 90th pct. (target 0.10)
JB — judgmental	2.866	2.510	0.641	0.359
TS — statistical	0.932	0.701	0.097	0.124
BDA — big data	0.812	0.595	0.099	0.098

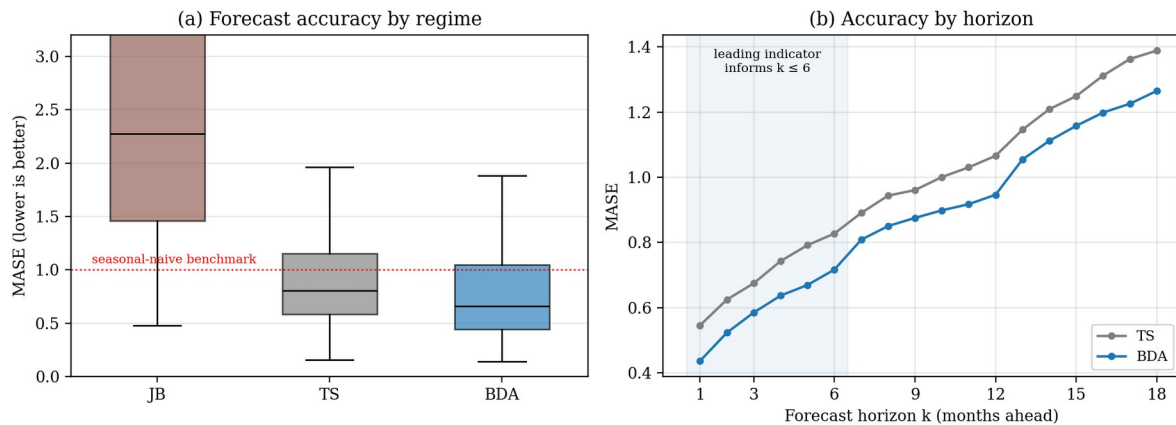


Figure 2. Forecast performance: (a) distribution of MASE by regime; (b) MASE by forecast horizon.

4.2. Planning performance

Table 6 reports realised performance over the 60-month planning window. Judgmental planning was the most expensive regime (0.3008 planning cost per unit) and the least reliable (2.89% of demand unmet). Quarterly statistical planning reduced planning cost by 21.0% relative to judgmental planning. The full BDA regime reduced it by a further 11.9% relative to TS, reaching 0.2093, and achieved the highest service level among feasible regimes (99.37% fill rate). Relative to judgmental planning, the BDA regime reduced planning cost by 30.4% and unmet demand from 2.89% to 0.63%. The Friedman test rejected equality across the six feasible regimes ($\chi^2 = 497.2$, $p < 0.001$), and all pairwise comparisons discussed below are significant at the 1% level unless stated otherwise.

Table 6. Realised planning performance (200 series $\times$ 60 months). Planning cost excludes standard production cost. Reduction is relative to TS; gap captured is the share of the JB-to-PI difference.

Regime	Planning cost per unit (s.d.)	Total cost per unit	Reduction vs. TS (%)	Gap captured (%)	Unmet demand (%)	Utilisation	Capacity churn	Nervousness
JB	0.3008 (0.216)	1.2678	-26.6	0.0	2.89	0.922	0.96	0.350
TS	0.2376 (0.135)	1.2193	0.0	48.8	1.24	0.991	0.91	0.071
TS-S	0.2339 (0.147)	1.2178	1.6	51.6	1.01	0.955	0.55	0.054
BDA-Q	0.2068 (0.088)	1.1939	13.0	72.6	0.69	0.969	0.40	0.032
BDA-D	0.2271 (0.094)	1.2141	4.4 (n.s.)	56.9	0.75	0.985	1.07	0.027
BDA	0.2093 (0.089)	1.1971	11.9	70.7	0.63	0.967	0.54	0.015
PI (bound)	0.1713 (0.017)	1.1675	27.9	100.0	0.02	1.002	0.31	0.005

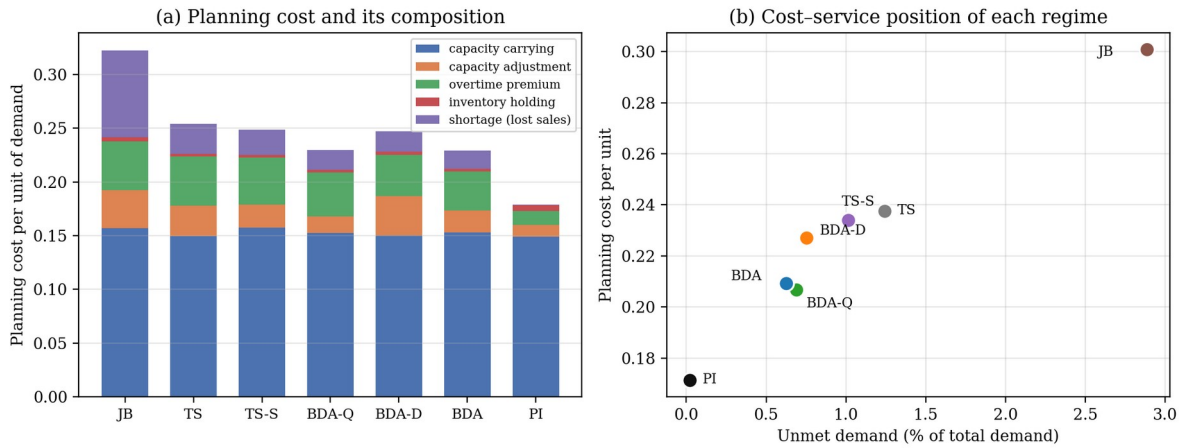


Figure 3. Planning performance: (a) planning cost and its composition; (b) cost–service position of each regime.

Figure 3(a) shows where the savings arise. Judgmental planning incurs a large shortage cost because its optimism bias and annual revision leave capacity misaligned with demand. Statistical planning removes most of that shortage cost but carries more capacity and adjusts it more often. The big-data regimes reduce shortage cost further while simultaneously lowering capacity adjustment cost, because more accurate distributional forecasts allow capacity to be positioned correctly rather than corrected repeatedly. Figure 3(b) shows that the BDA regimes dominate in both dimensions rather than trading cost against service.

4.3. Decomposing the value of the analytics upgrade

The distance between judgmental planning and perfect information (0.1295 per unit) is the improvement achievable by better information under this cost structure. Table 7 and Figure 4 decompose it into the components of the analytics upgrade. Formalising the forecast, that is moving from judgment to a quarterly statistical model, contributes 48.8% of the achievable gap and is by far the single largest step.

Adding an explicit predictive distribution to that statistical forecast contributes a further 2.8%. Improving forecast quality through cross-series learning and external data contributes 21.0%. Increasing the revision frequency from quarterly to monthly contributes –1.9%: it makes performance slightly but significantly worse ($p = 0.023$) while increasing capacity churn from 0.40 to 0.54. The residual 29.3% is irreducible under genuine uncertainty and would require information the firm cannot have.

Table 7. Decomposition of the planning-cost gap between judgmental planning and perfect information (200 paired series; Wilcoxon signed-rank tests).

Component of the upgrade	Contrast	Δ planning cost per unit	Share of achievable gap (%)	p-value
Formalising the forecast	JB $\rightarrow$ TS	–0.0632	48.8	2.9×10^{-19}
Representing uncertainty (statistical)	TS $\rightarrow$ TS-S	–0.0037	2.8	2.5×10^{-13}
Big-data forecast quality	TS-S $\rightarrow$ BDA-Q	–0.0272	21.0	9.8×10^{-11}
Increasing revision frequency	BDA-Q $\rightarrow$ BDA	+0.0025	–1.9	2.3×10^{-2}
Memo: representing uncertainty at monthly revision	BDA-D $\rightarrow$ BDA	–0.0178	13.8	1.6×10^{-29}
Irreducible uncertainty	BDA $\rightarrow$ PI	–0.0379	29.3	1.6×10^{-34}

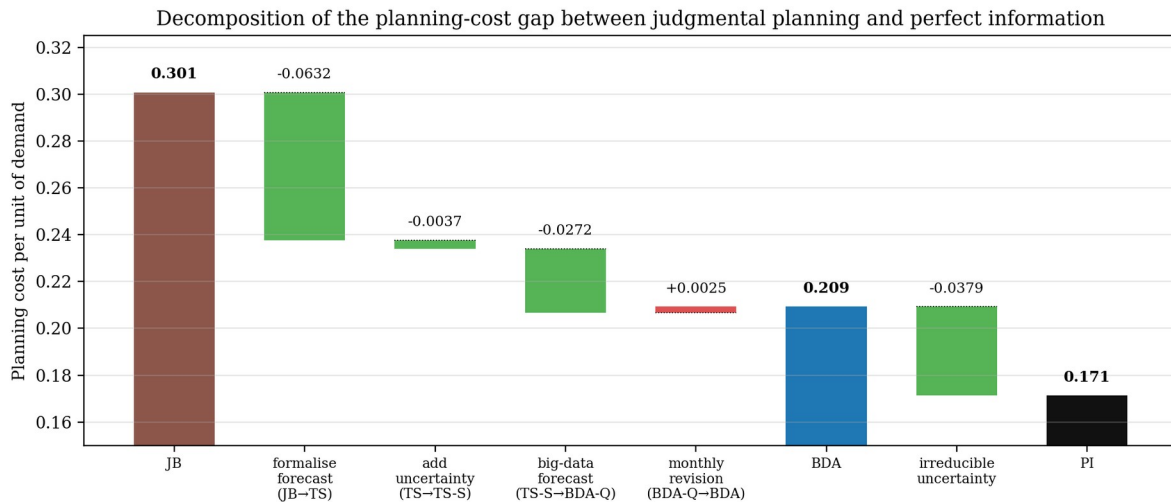


Figure 4. Decomposition of the planning-cost gap between judgmental planning and perfect information. Green bars reduce cost; the red bar increases it.

Two comparisons in Table 7 deserve emphasis. First, the value of representing uncertainty depends on the quality of the forecast it accompanies: it is small when added to the statistical forecast (2.8%) but large when added to the big-data forecast at monthly revision (13.8%). A sharper predictive distribution is worth more because the plan can act on it. Second, and more strikingly, the BDA-D regime, which uses the superior big-data forecast but collapses it to a point estimate, was statistically indistinguishable from simple quarterly statistical planning (0.2271 versus 0.2376; Wilcoxon $p = 0.94$) and had the highest capacity churn of any regime (1.07). Better prediction delivered no benefit when the decision model was not told how uncertain that prediction was.

4.4. Contingencies

Figure 5 and Table 8 report how the value of BDA depends on the operating context. The advantage over statistical planning increased monotonically with demand volatility, from 4.8% in the least volatile tercile to 17.2% in the most volatile. It also increased with the shortage penalty, from 9.2% when unmet demand costs 1.5 times unit production cost to 17.8% when it costs 6.0 times, because accurate distributional information is most valuable when the asymmetry between excess capacity and lost sales is large. The advantage decreased when the capacity lead time was long: at $L = 12$ months it fell to 9.2%, since commitments must then be made further ahead than the horizon over which the external signal and recent data are informative. Across all configurations, the BDA regime captured between 74.2% and 85.9% of the achievable gap.

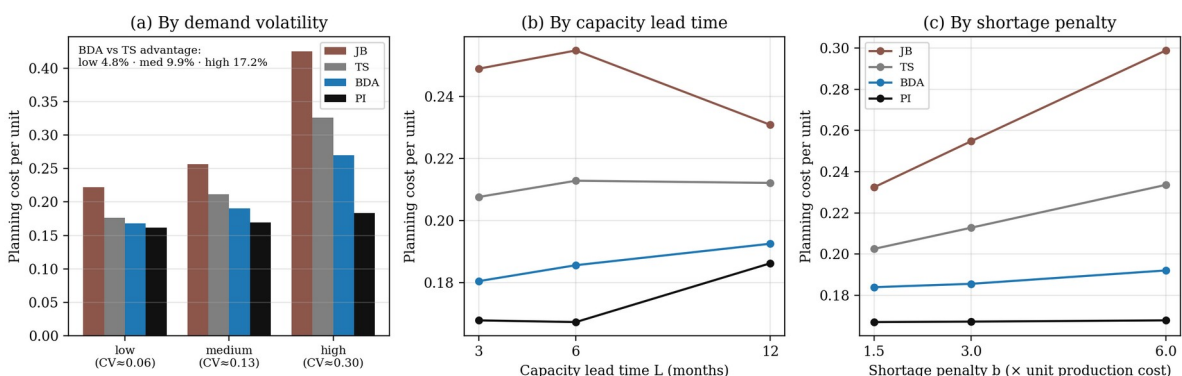


Figure 5. Contingency analysis: planning cost by (a) demand-volatility tercile, (b) capacity lead time, and (c) shortage penalty.

Table 8. Contingency analysis. Panel A: by demand-volatility tercile (200 series, base parameters). Panel B: by structural parameters (100 series).

Condition	JB	TS	BDA	PI	BDA vs. TS (%)	Gap captured (%)
Panel A — demand volatility						
Low ($CV \approx 0.06$)	0.2223	0.1765	0.1681	0.1614	4.8	89.0
Medium ($CV \approx 0.13$)	0.2564	0.2113	0.1904	0.1692	9.9	75.7
High ($CV \approx 0.30$)	0.4257	0.3264	0.2702	0.1836	17.2	64.2
Panel B — structural parameters						
Lead time $L = 3$ months	0.2490	0.2076	0.1804	0.1678	13.1	84.5
Lead time $L = 6$ months (base)	0.2548	0.2128	0.1856	0.1672	12.8	79.0
Lead time $L = 12$ months	0.2309	0.2121	0.1925	0.1862	9.2	85.9
Shortage penalty $b = 1.5$	0.2325	0.2025	0.1839	0.1670	9.2	74.2
Shortage penalty $b = 6.0$	0.2988	0.2336	0.1921	0.1678	17.8	81.5

4.5. The contribution of external data

Setting the leading indicator to pure noise ($\rho = 0$) isolates the contribution of external data from that of cross-series learning. Without the external signal, the global model's accuracy deteriorated by 2.1% overall and 3.3% over the first six months, and planning cost rose by 1.1% for the full BDA regime (all $p < 0.01$, 100 series). The statistical regime, which does not use the signal, was unaffected, confirming the internal validity of the comparison. The external signal is therefore genuinely useful but contributes only a small part of the BDA advantage: even without it, the global model still outperformed the per-series statistical model substantially. At the signal quality examined here, the dominant source of big-data value in this setting is learning across many related series rather than the external indicator itself.

Table 9. Ablation of the external leading indicator (100 series).

Measure	With signal ($\rho = 0.5$)	Without signal ($\rho = 0$)	Difference (%)	p-value
BDA forecast MASE (all horizons)	0.702	0.717	-2.1	2.2×10^{-9}
BDA forecast MASE (horizons 1–6)	0.496	0.513	-3.3	4.0×10^{-5}
BDA planning cost per unit	0.1856	0.1876	-1.1	1.1×10^{-3}
BDA-Q planning cost per unit	0.1853	0.1867	-0.7	2.3×10^{-3}
TS planning cost per unit (control)	0.2128	0.2128	0.0	n.a.

5. Discussion

5.1. Interpretation

On RQ1, the big-data regime improved planning performance substantially, reducing planning cost by 30.4% relative to judgmental planning and 11.9% relative to quarterly statistical planning while improving service, and capturing roughly 71% of the improvement attainable with perfect information. On RQ2, the improvement is unevenly distributed. Nearly half of the achievable gap is captured simply by replacing judgment with a formal statistical forecast, a step requiring no big data at all. Big-data forecast quality adds a further fifth, and representing uncertainty adds a substantial share once forecasts are sharp enough to be worth acting on. Increasing revision frequency, the component most closely associated with the rhetoric of real-time analytics, contributed nothing and marginally degraded performance. On RQ3, the value of BDA rises with demand volatility and with the cost of shortages, and falls when capacity commitments must be made further ahead than the data can see.

The finding that more frequent replanning does not help warrants explanation. With a six-month acquisition lead time, a decision taken this month cannot be undone before it takes effect. Revising monthly therefore reacts to noise in the most recent observations, generating capacity adjustments that are later reversed; capacity churn rose by a third and nervousness measures deteriorated for the deterministic monthly regime. This is the planning analogue of the nervousness long recognised in rolling-horizon MRP systems [14], [13]. It suggests that the marginal value of data velocity is bounded above by the responsiveness of the underlying physical system, an argument that the analytics literature, focused on capability constructs, has had little means to test [5], [8].

5.2. Theoretical implications

The results support and refine the prescriptive-analytics argument that data creates value only through the decision it informs [16]. The BDA-D regime provides a clean demonstration: a forecast that was 12.9% more accurate produced no significant cost improvement when its uncertainty was discarded. This has a methodological consequence for research design. Studies that evaluate industrial analytics by prediction error alone can rank methods in an order that does not correspond to their operational value, and studies that measure analytics capability as a composite construct cannot detect that a component with high perceived salience, such as real-time dashboards, may have low or negative marginal return. Decomposition designs of the kind used here offer a complementary route to the survey tradition [5], [6], [7], [8], and connect it to the operations literature on planning under uncertainty [10], [15].

The contingency results also qualify the resource-based reading of analytics. The same analytics system generated between 4.8% and 17.8% improvement depending on volatility and shortage economics, so the operational return is a property of the decision context as much as of the capability. This is consistent with the moderating role of environmental conditions reported in the capability literature [8] and gives it an operational mechanism.

5.3. Managerial implications

Four recommendations follow. First, firms still planning judgmentally should prioritise formalising the forecasting process, since that step alone captured almost half of the achievable improvement and requires no new data infrastructure. Second, analytics investment should be directed at decision quality before data velocity: carrying a predictive distribution into a cost-aware planning model was consistently valuable, whereas shortening the replanning cycle was not. Third, the review interval should be matched to the lead time of the decision rather than to the arrival rate of data; where

commitments have long lead times, more frequent revision mainly generates instability. Fourth, the expected return should be assessed against the firm's own context: analytics pays most where demand is volatile and shortages are expensive, and least where demand is stable or capacity decisions must be taken far in advance. A firm in the latter position may reasonably conclude that a well-run quarterly statistical process is sufficient.

5.4. Limitations

Several limitations should be noted. Demand paths are real, but the planning environment is a single-product, single-stage capacity system; multi-product interactions, supplier constraints and multi-echelon effects are not represented. Costs are normalised and the shortage penalty is a scenario assumption, although the sensitivity analysis spans a wide range. The judgmental regime is a stylised model of biased planning rather than an observed human process; its parameters were set from documented behavioural patterns [25], [26] and not estimated from data, so the size of the judgmental gap should be read as indicative. The external leading indicator is synthetic, which allows signal quality to be controlled but does not reproduce the noise structure of real indicators [17]; the ablation quantifies its contribution at one signal level only. Finally, the organisational costs of operating each regime, including data infrastructure, model maintenance and analyst time, are not modelled, so the results describe gross rather than net returns to analytics.

6. Conclusion and Future Research

This study asked whether bigger data makes better plans, and which part of an analytics upgrade delivers the improvement. Using 200 real industrial demand series and an identical stochastic planning and execution environment, it found that a full big-data planning regime reduced planning cost by 30.4% relative to judgmental planning and by 11.9% relative to quarterly statistical planning, capturing about 71% of the improvement achievable with perfect information. Decomposition showed that formalising the forecast contributed 48.8% of that achievable improvement, big-data forecast quality 21.0%, and explicit uncertainty representation up to 13.8%, while increasing revision frequency from quarterly to monthly contributed nothing and slightly increased cost and plan instability. A more accurate forecast used deterministically performed no better than simple quarterly statistical planning. The value of analytics increased with demand volatility and shortage cost and decreased with capacity lead time.

Future research should extend the decomposition to multi-product and multi-echelon settings where analytics may have different leverage; replace the synthetic leading indicator with real external data to examine how signal quality and timeliness interact with lead time; incorporate the cost of operating an analytics system to move from gross to net return; test the judgmental baseline against forecasts produced by human planners; and examine whether the negative return to revision frequency persists when capacity is more flexible, which would delineate the conditions under which real-time planning becomes worthwhile.

Declarations

Data availability. The demand data are derived from the publicly available M4 competition dataset (<https://github.com/Mcompetitions/M4-methods>); the identifiers of the 200 series used are provided in the supplementary material. Code availability. All simulation and analysis code, with fixed random seeds, is provided in the supplementary material. Funding. [To be completed by the authors.] Conflict of interest. The author declares no conflict of interest. Author contributions. [To be completed.]

Declaration of generative AI use. [Authors: most publishers require disclosure of generative AI tools used in preparing a manuscript. Please complete this statement according to the target journal's policy.]

References

- [1] L. D. Xu, E. L. Xu, and L. Li, "Industry 4.0: State of the art and future trends," *Int. J. Prod. Res.*, vol. 56, no. 8, pp. 2941–2962, 2018, doi: 10.1080/00207543.2018.1444806.
- [2] J. Lee, H.-A. Kao, and S. Yang, "Service innovation and smart analytics for Industry 4.0 and big data environment," *Procedia CIRP*, vol. 16, pp. 3–8, 2014, doi: 10.1016/j.procir.2014.02.001.
- [3] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and Industry 5.0—Inception, conception and perception," *J. Manuf. Syst.*, vol. 61, pp. 530–535, Oct. 2021, doi: 10.1016/j.jmsy.2021.10.006.
- [4] T.-M. Choi, S. W. Wallace, and Y. Wang, "Big data analytics in operations management," *Prod. Oper. Manage.*, vol. 27, no. 10, pp. 1868–1883, 2018, doi: 10.1111/poms.12838.
- [5] S. F. Wamba, A. Gunasekaran, S. Akter, S. J. Ren, R. Dubey, and S. J. Childe, "Big data analytics and firm performance: Effects of dynamic capabilities," *J. Bus. Res.*, vol. 70, pp. 356–365, Jan. 2017, doi: 10.1016/j.jbusres.2016.08.009.
- [6] A. Gunasekaran, T. Papadopoulos, R. Dubey, S. F. Wamba, S. J. Childe, B. Hazen, and S. Akter, "Big data and predictive analytics for supply chain and organizational performance," *J. Bus. Res.*, vol. 70, pp. 308–317, Jan. 2017, doi: 10.1016/j.jbusres.2016.08.004.
- [7] B. Chae, D. Olson, and C. Sheu, "The impact of supply chain analytics on operational performance: A resource-based view," *Int. J. Prod. Res.*, vol. 52, no. 16, pp. 4695–4710, 2014, doi: 10.1080/00207543.2013.861616.
- [8] P. Mikalef, M. Boura, G. Lekakos, and J. Krogstie, "Big data analytics capabilities and innovation: The mediating role of dynamic capabilities and moderating effect of the environment," *Br. J. Manage.*, vol. 30, no. 2, pp. 272–298, 2019, doi: 10.1111/1467-8551.12343.
- [9] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 Competition: 100,000 time series and 61 forecasting methods," *Int. J. Forecast.*, vol. 36, no. 1, pp. 54–74, 2020, doi: 10.1016/j.ijforecast.2019.04.014.
- [10] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, 2nd ed. New York, NY, USA: Springer, 2011.
- [11] E. Brynjolfsson, D. Rock, and C. Syverson, "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," in *The Economics of Artificial Intelligence: An Agenda*, A. Agrawal, J. Gans, and A. Goldfarb, Eds. Chicago, IL, USA: Univ. of Chicago Press, 2019, pp. 23–57.
- [12] Y. K. Dwivedi et al., "Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *Int. J. Inf. Manage.*, vol. 57, Art. no. 101994, Apr. 2021, doi: 10.1016/j.ijinfomgt.2019.08.002.
- [13] W. J. Hopp and M. L. Spearman, *Factory Physics*, 3rd ed. Long Grove, IL, USA: Waveland Press, 2011.
- [14] E. A. Silver, D. F. Pyke, and D. J. Thomas, *Inventory and Production Management in Supply Chains*, 4th ed. Boca Raton, FL, USA: CRC Press, 2016.
- [15] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, 2nd ed. Philadelphia, PA, USA: SIAM, 2014.
- [16] D. Bertsimas and N. Kallus, "From predictive to prescriptive analytics," *Manage. Sci.*, vol. 66, no. 3, pp. 1025–1044, Mar. 2020, doi: 10.1287/mnsc.2018.3253.
- [17] Y. R. Sagaert, E.-H. Aghezzaf, N. Kourentzes, and B. Desmet, "Tactical sales forecasting using a very large set of macroeconomic indicators," *Eur. J. Oper. Res.*, vol. 264, no. 2, pp. 558–569, 2018, doi: 10.1016/j.ejor.2017.06.054.

- [18] A. Goli, H. K. Zare, R. Tavakkoli-Moghaddam, and A. Sadeghieh, "Hybrid artificial intelligence and robust optimization for a multi-objective product portfolio problem. Case study: The dairy products industry," *Comput. Ind. Eng.*, vol. 137, Art. no. 106090, 2019.
- [19] A. Goli, H. Khademi-Zare, R. Tavakkoli-Moghaddam, A. Sadeghieh, M. Sasanian, and R. Malekalipour Kordestanizadeh, "An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: A case study," *Network: Comput. Neural Syst.*, vol. 32, no. 1, pp. 1–35, 2021.
- [20] A. Goli, A. M. Golmohammadi, and S. A. Edalatpanah, "Application of artificial intelligence in forecasting the demand for supply chains considering Industry 4.0," in *A Roadmap for Enabling Industry 4.0 by Artificial Intelligence*. Hoboken, NJ, USA: Wiley, 2022, pp. 43–55.
- [21] A. Goli, "Efficient optimization of robust project scheduling for Industry 4.0: A hybrid approach based on machine learning and meta-heuristic algorithms," *Int. J. Prod. Econ.*, vol. 278, Art. no. 109427, 2024.
- [22] E. B. Tirkolaee, A. Goli, and G.-W. Weber, "Fuzzy mathematical programming and self-adaptive artificial fish swarm algorithm for just-in-time energy-aware flow shop scheduling problem with outsourcing option," *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 11, pp. 2772–2783, Nov. 2020.
- [23] P. Ghasemi, A. Goli, F. Goodarzian, and J. F. Ehmke, "Simulation-based genetic algorithm for optimizing a municipal cooperative waste supply chain in a pandemic," *Eng. Appl. Artif. Intell.*, vol. 139, Art. no. 109478, 2025.
- [24] A. M. Golmohammadi, A. Goli, N. Jahanbakhsh-Javid, and H. Farughi, "Simultaneous consideration of time and cost impacts of machine failures on cellular manufacturing systems," *Eng. Appl. Artif. Intell.*, vol. 134, Art. no. 108480, 2024.
- [25] A. Tversky and D. Kahneman, "Judgment under uncertainty: Heuristics and biases," *Science*, vol. 185, no. 4157, pp. 1124–1131, 1974.
- [26] M. Lawrence, P. Goodwin, M. O'Connor, and D. Önkal, "Judgmental forecasting: A review of progress over the last 25 years," *Int. J. Forecast.*, vol. 22, no. 3, pp. 493–518, 2006, doi: 10.1016/j.ijforecast.2006.03.007.
- [27] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006, doi: 10.1016/j.ijforecast.2006.03.001.
- [28] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021.
- [29] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [30] R. Koenker and G. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.
- [31] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. Amer. Statist. Assoc.*, vol. 102, no. 477, pp. 359–378, 2007.
- [32] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, 2006.
- [33] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.