JURNAL LOCUS: Penelitian & Pengabdian
Volume 4 No 10 Oktober 2025
E-ISSN 2829-7334| P-ISSN 2829-5439
Hompage: https://locus.rivierapublishing.id/index.php/jl

Strategic Operations under Uncertainty: Mitigating Shutdown Risks and Operator Error through AI-Driven Decision Support

Anton Santoso, Arviansyah
Universitas Indonesia
Email: anton_santoso@outlook.com

Abstract:
In complex industrial environments such as integrated refinery and petrochemical systems, operational mistakes and unexpected unit shutdowns pose significant strategic challenges. These incidents not only disrupt production continuity but also create substantial financial and safety risks. This study examines the strategic role of artificial intelligence (AI) in enhancing operations management by supporting decision-making during abnormal conditions. A real-case-based simulation was conducted using dynamic models of an integrated plant configuration, with experienced operators performing under both AI-assisted and conventional manual scenarios. Four shutdown scenarios, both detected and undetected, were simulated to evaluate adaptability, multitasking capacity, and resilience. Performance was measured using five key indicators: time to production target, product yield, operational deviations, utility efficiency, and inventory utilization. Statistical analysis employing the Mann–Whitney U test and Bayesian inference revealed that AI assistance significantly enhanced operator performance across all dimensions, particularly by reducing errors and optimizing response times. This research contributes to the ongoing transformation of industrial decision-making by highlighting AI's potential to improve resilience, reduce cognitive load, and mitigate risks under operational uncertainty. Although limitations related to sample size and scenario diversity are acknowledged, they do not undermine the reliability or practical applicability of the findings for managerial decision support systems.

Keywords: Operations Management; Artificial Intelligence (AI) Decision Support; Industrial Shutdown Mitigation; Operator Performance and Error Reduction; Resilience in Process Industries

Corresponding: Anton Santoso
E-mail: anton_santoso@outlook.com

INTRODUCTION

In integrated refinery and petrochemical operations, managing capacity across interconnected units is a highly complex task that demands both strategic oversight and operational agility (DOZIE, JOHNSON, ANDREW, & JUSTINE, 2023; Ojuekaiye, 2024). At the local level, control of individual processing units is typically handled by lower-level operators using established procedures and automation systems (Sharma, 2016). However, the responsibility for managing overall production capacity, particularly across multiple units and processes, falls on senior engineers and operations managers (Rao, 2023). These individuals are responsible not only for maintaining system balance and efficiency but also for coordinating with other key functions such as procurement, planning, and marketing (Kerzner, 2025). This cross-functional coordination becomes even more critical in integrated facilities, where upstream and downstream processes are tightly interlinked (Brewer, Ashenbaum, & Blair, 2019; Swink & Schoenherr, 2015).

The industrial landscape faces mounting challenges due to increasing operational complexity, regulatory demands, and economic pressures. Recent industry data from 2024 indicates that unplanned shutdowns cost the global refining industry approximately $50 billion annually, with 40% of these incidents attributed to operator error or delayed responses to critical situations. In petrochemical operations specifically, even minor disruptions can cascade through interconnected systems, leading to production losses exceeding $2 million per day for large-scale facilities. These statistics underscore the critical need for advanced decision-support systems that can enhance operator performance during high-stress scenarios.

The scope of work in such integrated systems is vast and often non-linear. A small operational mistake, such as a delayed response to a unit upset or a misjudged feedstock adjustment, can cascade through the entire system, causing significant production losses, product off-spec events, or increased utility consumption. In some cases, these mistakes may be absorbed as tolerable inefficiencies, especially when market conditions provide an operational buffer or when energy costs are low (Hirsch, Parag, & Guerrero, 2018; Medina, Ana, & González, 2022; Meegahapola, Mancarella, Flynn, & Moreno, 2021). Current market conditions, however, have intensified these challenges significantly. Energy costs have increased by 35% since 2023, margin pressures have tightened due to volatile crude oil prices ranging from $70–95 per barrel, and new environmental regulations mandate a 20% reduction in emissions by 2030 (Bhattarai & Yousef, 2025; Kaljunen, 2024). These factors have eliminated the operational buffer that previously allowed for inefficiencies, making precision in decision-making critical for maintaining competitiveness and compliance.

The urgency for implementing AI-driven solutions has been further amplified by recent technological advances and adoption trends. According to 2024 industry surveys, 65% of major oil and gas companies have initiated AI pilot programs, with 23% reporting significant operational improvements within the first year of implementation. However, most implementations focus on predictive maintenance and basic optimization, leaving a significant gap in real-time operational decision support during crisis situations.

In this context, the application of artificial intelligence (AI) has emerged as a promising solution (Jiang et al., 2021; Xu et al., 2021). Numerous studies have highlighted the benefits of AI in industrial environments, particularly for optimization, anomaly detection, predictive maintenance, and advanced decision support (Rojas, Peña, & Garcia, 2025). Unlike traditional rule-based systems, AI can learn from historical data, adapt to new conditions, and provide real-time insights to support more informed and faster decision-making (Tien, 2017). Despite this growing body of research, the application of AI in the specific context of operational integration, particularly within complex, real-world refinery and petrochemical systems, remains relatively underexplored (Bigdeli & Delshad, 2024).

This study seeks to bridge that gap by evaluating how AI can assist human operators in managing operational integration more effectively. We focus on practical scenarios that reflect the challenges faced by senior operators and operations managers, such as sudden unit trips, planned shutdowns, or feedstock limitations. Through a series of simulations, we replicate real-world disturbances and measure how decision-making performance changes with and without AI support.

To ensure relevance and realism, our scenario development was based on input from experienced practitioners through a series of structured interviews. We then designed a controlled testing environment in which two groups of operators, each with comparable levels of experience, managed the same simulation cases. One group operated the system using conventional methods, relying solely on their experience and system feedback, while the other group was provided with AI-assisted decision-support tools. The data collected from these simulations were analyzed to evaluate performance differences in terms of response time, decision accuracy, system stability, and overall operational impact.

Despite growing research interest, the application of AI in real-time operational decision support for integrated refinery and petrochemical systems remains significantly underexplored, particularly in crisis management scenarios where human operators must make rapid, high-stakes decisions under uncertainty.

This research addresses several critical gaps in current knowledge. First, while existing studies primarily focus on AI for routine optimization tasks, there is limited empirical evidence regarding AI's effectiveness in supporting human operators during emergency shutdown scenarios. Second, most AI implementations in process industries operate as standalone systems, whereas this study investigates human-AI collaboration in which operators remain actively engaged in decision-making. Third, previous research has typically employed theoretical models or laboratory simulations, while this study utilizes realistic plant configurations and scenarios based on actual industrial experience.

The primary objective of this research is to evaluate the effectiveness of AI-assisted decision support in enhancing operator performance during critical operational disruptions in integrated refinery and petrochemical systems. Specific aims include: (1) quantifying performance improvements across multiple operational dimensions when AI support is available, (2) assessing the impact of AI assistance on error reduction and response-time optimization, (3) determining the effectiveness of AI support across different types of shutdown scenarios (detected vs. undetected), and (4) providing empirical evidence to support strategic decisions regarding AI implementation in industrial control environments.

The practical benefits of this research extend beyond academic contribution. For industry practitioners, the findings provide quantitative evidence to justify investments in AI-driven decision-support systems. For operators and plant managers, the research demonstrates how AI can augment rather than replace human expertise, potentially reducing cognitive load during stressful situations while maintaining human oversight and control. For technology developers, the study offers insights into effective human-AI interface design for critical industrial applications.

RESEARCH METHOD

This study employs a quantitative experimental research design to compare operational performance between conventional manual control and artificial intelligence (AI)-assisted decision-making in an integrated refinery and petrochemical complex. The objective is to assess the effectiveness of AI in managing operational disruptions, particularly in dynamic and high-risk scenarios.

A pre-experimental simulation-based approach is used to represent actual operational conditions. Simulations are constructed based on real plant configurations and typical operational challenges. Operators with substantial experience are randomly selected from a qualified pool to participate. Each operator is required to complete the same set of scenarios twice: once using manual decision-making (without AI support) and once using AI-assisted systems.

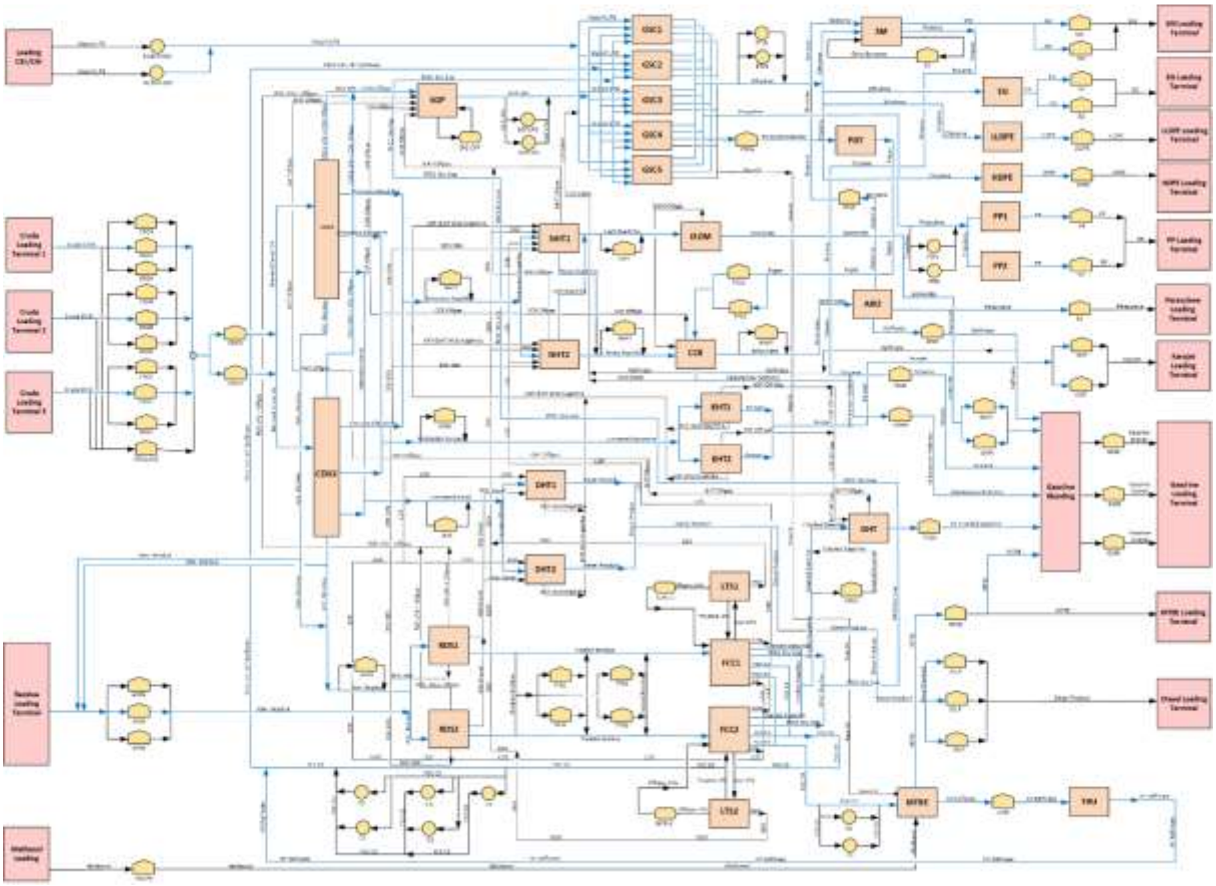

Figure 1. The configuration of Integration Refinery and Petrochemical used for the case study.

Plant Configuration and System Description

The simulation models a large-scale integrated refinery and petrochemical complex with a daily processing capacity of 400,000 barrels. The facility processes a fixed blend of three imported crude oils (A, B, and C) through two Crude Distillation Units (CDU1 and CDU2) and produces a range of fuel and petrochemical products, including gasoline, diesel, kerosene, LLDPE, HDPE, PP, styrene monomer (SM), and ethylene glycol (EG).

The process flow involves multiple downstream units such as:

- Naphtha Hydrotreating, ISOM, and CCR for gasoline blending and aromatics production,
- Hydrotreating units for diesel and kerosene,
- Resid Fluid Catalytic Cracking (RFCC1 and RFCC2),
- Steam Crackers (GSC1–GSC5) for ethylene and propylene production,
- Pygas Treatment for recycling aromatic intermediates,
- Utility systems and storage tanks for feedstock blending and product buffering.
- This complexity presents ideal conditions for testing operator responsiveness, especially during unit failures.

Simulation Scenario Design

The simulation scenario includes four dynamic operational disruptions (individual unit shut-downs) based on real cases gathered through interviews with senior practitioners, as summarized in Table 1. The scenario simulates different types of unit failure, representing both detected (with early warning) and undetected (sudden) disturbances.

Operators must perform mitigation or recovery actions in advance, depending on whether the failure is detected. At the same time, they are responsible for routine operations and meeting production targets. The simulation evaluates not only technical decision-making but also multitasking under pressure.

Performance Indicators

The effectiveness of operator performance under both manual and AI-assisted modes is assessed using the following metrics:

- Time to Target: Measures the duration required to reach production targets after a disturbance, indicating response speed and planning effectiveness.
- Product Yield: Assesses the conversion efficiency of feedstock into key products, reflecting material utilization efficiency.
- Operational Deviations: Counts the number of procedural errors made during simulation, including missed steps, misrouted flows, or late adjustments.
- Utility Efficiency: Calculates the ratio of critical utilities (e.g., electricity, steam, cooling water) used relative to output, assessing resource management.
- Inventory Utilisation: This evaluates the amount of buffer inventory used to mitigate disruptions, revealing the operator's ability to balance production continuity with limited storage.

Simulation Execution and Data Collection

The process simulation was built using actual technical process data and validated through a pilot test with operators not included in the final test group. After refinements, the finalized simulation was used for the main experiment. Each operator completed all designed scenarios under control conditions, either manually (without AI) or AI-assisted. To ensure consistency, simulations were conducted individually to prevent cross-contamination of decisions or peer influence. The AI system employed reinforcement learning algorithms with real-time influence diagrams to support decision-making. Simulation outputs were logged automatically. Performance data for each operator in each scenario were extracted directly from the system and statistically analyzed to determine significant differences in outcomes between the two modes.

Statistical Analysis Approach

Given the study design comparing two independent groups (AI-assisted vs. manual operation) across five distinct performance metrics, a careful statistical treatment is necessary to avoid inflated Type I error rates due to multiple comparisons. To mitigate this, we apply the Bonferroni correction, which adjusts the significance threshold by dividing the standard alpha level ($\alpha = 0.05$) by the number of independent tests conducted. With five parameters being analyzed, the revised threshold for statistical significance becomes $\alpha = 0.01$. This conservative approach ensures that any detected differences are less likely to be due to random variation.

The Mann–Whitney U test, a non-parametric alternative to the t-test, is chosen due to the small sample size and the ordinal nature of the performance data. This test does not assume a normal distribution and is suitable for detecting differences in central tendency based on ranked data. However, beyond the p-value, we place particular emphasis on the rank biserial correlation coefficient (r_s), which measures the strength and direction of the association between the two groups. This coefficient quantifies how consistently one group scores higher or lower than the other, providing an interpretable measure of effect size. According to conventional benchmarks, an r_s value around 0.1 indicates a small effect, 0.3 a moderate effect, and 0.5 or higher a strong effect.

Because of the limited number of participants (n = 3 per group), traditional significance testing may lack the power to detect meaningful differences. Therefore, we complement the analysis with a Bayesian independent-sample t-test, which offers a more nuanced assessment of the evidence. Instead of yielding a binary decision (reject or retain the null hypothesis), the Bayesian framework estimates the Bayes Factor (BF_{10}), which represents how much more likely the observed data are under the alternative hypothesis than under the null.

Bayes Factors are interpreted along a continuum: values between 1 and 3 suggest weak or anecdotal evidence in favor of a difference, values between 3 and 10 indicate moderate support, values between 10 and 30 suggest strong evidence, and anything above 30 is typically considered very strong or decisive. On the other hand, values below 1 favor the null hypothesis, implying that the data better support the absence of a difference.

Importantly, the Bayesian approach remains valid even with small samples, although the interpretation becomes more sensitive to the choice of prior distributions. In this context, the Bayes Factor serves as a valuable supplement to classical testing, particularly when p-values fall just above traditional significance thresholds, but the practical implications of the observed differences may still be meaningful. Together, these statistical tools provide a robust foundation for evaluating the impact of AI support on operator performance across complex operational scenarios.

Tabel 1. Failure Scenario in the Studi Case During The Simulation

Scenario	Unit	Type of Failure	Notification	Failure Starting	Duration
1	GSC1	Detected	Day 4	Day 8	2 days
2	KHT1	Undetected	-	Day15	3 days
3	CDU2	Detected	Day16	Day 20	4 days
4	RFCC2	Undetected	-	Day 22	6 days

RESULTS

The performance outcomes of operator-only control (without AI assistance) and AI-assisted operation (with operator involvement) were compared across five critical variables using the Mann–Whitney U test. Due to the limited sample size (n = 3 per group), p-values were interpreted cautiously, and the rank biserial correlation was used to assess effect sizes. Additionally, the Bonferroni correction (adjusted $\alpha = 0.01$) was considered to account for multiple testing.

Operators working with AI support reached production targets considerably faster than those managing operations manually. In the manual group, the time to target ranged from 0.508 to 0.709, while the AI-assisted group showed values from -0.119 to 0.174. The Mann–Whitney U test yielded a p-value of 0.05 and a maximum effect size ($r_s = 1.00$), indicating a very strong difference in performance favoring AI-assisted operations.

Product yield improved notably with AI assistance. The manual group achieved yields between 0.650 and 0.717, while the AI-assisted group reached 0.759 to 0.826. The difference was statistically significant (p = 0.05) with a large effect size (1.00), signifying a consistent and meaningful improvement in production efficiency due to AI support.

Table 2. Comparison of simulation results in terms of achieving production targets

	Target	Without AI			With AI		
Variables	(day)	Op. 1	Op. 2	Op. 3	Op. 4	Op. 5	Op. 6
Time to achieve kerosene 4000 KB	40	45	46	43	40	41	42
Time to achieve diesel 3400 KB	31	35	36	35	30	31	32
Time to achieve kerosene 1000 kB	30	34	35	33	30	30	31
Time to achieve Paraxylene 1000KB	33	36	36	36	32	33	34
Time to achieve PE 64000 tons	35	37	37	37	34	35	36
Time to achieve PP 68000 tons	36	38	39	38	35	36	36
$\sum(\Delta T/T)$		0,591	0,709	0,508	-0,119	0,025	0,174

Table 3. Comparison of simulation results in terms of production yield

	Without AI			With AI		
Variables	**Op. 1**	**Op. 2**	**Op. 3**	**Op. 4**	**Op. 5**	**Op. 6**
Time (days)	40	40	38	36	36	38
Inputs (Ton)	2.056.659,71	1.881.376,22	1.713.883,09	1.558.075,54	1.604.817,81	1.869.690,65
Output (Ton)	1.337.773,90	1.299.792,82	1.228.311,57	1.286.347,09	1.319.941,74	1.419.010,98
Yield	0,679	0,595	0,693	0,796	0,809	0,768

AI-assisted operations led to a dramatic reduction in operational deviations. Manual control resulted in 20 to 26 deviations, whereas the AI-assisted group showed only 0 to 4 deviations. With a p-value of 0.05 and an effect size of 1.00, the data reveal a substantial enhancement in operational discipline and procedural accuracy enabled by AI support.

The utility ratio significantly increased in AI-assisted operations. Manual operations recorded ratios between 0.779 and 0.827, while AI-assisted operations ranged from 0.924 to 0.931. This difference, statistically significant (p = 0.05) and with a very large effect size, suggests more efficient use of utilities such as energy and raw materials under AI-assisted control.

Inventory discrepancy was lower in the AI-assisted group, decreasing from 0.833–0.844 (manual) to 0.915–0.931 (AI-assisted). Despite the modest numerical change, the difference was statistically significant (p = 0.05) with a strong effect size (1.00), indicating improved material balance and inventory control when using AI.

Table 4: Comparison of simulation results in terms of total mistakes happening during the operation

	Without AI			With AI		
Variable	**Op. 1**	**Op. 2**	**Op. 3**	**Op. 4**	**Op. 5**	**Op. 6**
Mistakes	20	22	26	0	2	4

Table 5: Comparison of simulation results in terms of intermediate buffer tank utilization

	Target	Without AI			With AI		
Variables	**(day)**	**Op. 1**	**Op.2**	**Op.3**	**Op.4**	**Op.5**	**Op.6**
FG (Cons.)	Tons/hr	159,96	154,52	155,04	174,31	194,39	171,93
FG (Gen.)	Tons/hr	196,75	203,96	189,15	188,26	207,99	190,84
POWER (Cons.)	KW	472.702,73	456.615,60	458.158,47	515.125,05	574.442,48	515.950,05

	Target	Without AI			With AI		
Variables	**(day)**	**Op. 1**	**Op.2**	**Op.3**	**Op.4**	**Op.5**	**Op.6**
POWER (Gen.)	KW	595.605,43	611.864,90	563.534,92	561.486,30	620.397,88	577.864,06
HS (Cons.))	Tons/hr	1.175,63	1.136,38	1.140,22	1.281,97	1.429,59	1.224,22
HS (Gen.)	Tons/hr	1.504,80	1.454,56	1.413,87	1.371,71	1.543,96	1.346,64
MS (Cons.)	Tons/hr	650,10	688,16	630,55	708,92	769,93	709,50
MS (Gen.)	Tons/hr	773,62	853,32	725,14	758,55	831,53	773,36
BFW (Cons.)	Tons/hr	1.187,18	1.148,09	1.203,12	1.317,12	1.442,56	1.282,88
BFW (Gen.)	Tons/hr	1.780,76	1.412,15	1.455,78	1.396,15	1.557,96	1.398,33
Ratio FG	-	0,8130	0,7576	0,8197	0,9259	0,9346	0,9009
Ratio POWER	-	0,7937	0,7463	0,8130	0,9174	0,9259	0,8929
Ratio HS	-	0,7813	0,7813	0,8065	0,9346	0,9259	0,9091
Ratio MS	-	0,8403	0,8065	0,8696	0,9346	0,9259	0,9174
Average		0,6667	0,8130	0,8264	0,9434	0,9259	0,9174

*Cons = Consumption, Gen. = Generation

Table 6. Comparison of simulation results in terms of intermediate buffer tank utilization

	Without RTO			With RTO		
Variable	**Op.1**	**Op.2**	**Op.3**	**Op.4**	**Op.5**	**Op.6**
RV_{Total} - ΔRV	15,19	14,9875	15,0825	16,76	16,67	16,4625
RV_{total}	18	18	18	18	18	18
Ratio	0,8439	0,8326	0,8379	0,9311	0,9261	0,9146

Tabel 7. Mann-Whitney Method Independent Sample T-Test Result

		Without AI			With AI			Mann Whitney U Test		
	Variabel	**Op.1**	**Op.2**	**Op.3**	**Op.4**	**Op.5**	**Op.5**	**Hipotesis**	**p value**	**Effect Size**
1	Time to Target	0,591	0,709	0,508	-0,119	0,025	0,174	$Q_1 > Q_2$	0.05	1.00
2	Product Yield	0,650	0,691	0,717	0,826	0,822	0,759	$Q_1 < Q_2$	0.05	1.00
3	Operational Deviations	20	22	26	0	2	4	$Q_1 > Q_2$	0.05	1.00
4	Utility Ratio	0,779	0,781	0,827	0,931	0,928	0,924	$Q_1 < Q_2$	0.05	1.00
5	Inventory Discrepancy	0,844	0,833	0,838	0,931	0,926	0,915	$Q_1 < Q_2$	0.05	1.00

Tabel 8. Bayesian Independent Sample T-test Result

	Bayesian Independent Sample T-test				
	Variabel	**Hypothesis**	**p value**	**Effect Size**	**Remarks**
1	Time to Target	$Q_1 > Q_2$	15,485	$-7.204*10^{-5}$	Accepted
2	Product Yield	$Q_1 < Q_2$	7,306	$-8,149*10^{-4}$	Accepted
3	Operational Deviations	$Q_1 > Q_2$	64,098	$-1,143*10^{-6}$	Accepted
4	Utility Ratio	$Q_1 < Q_2$	41,805	$-3,688*10^{-7}$	Accepted
5	Inventory Discrepancy	$Q_1 < Q_2$	211,428	$-3,702*10^{-4}$	Accepted

DISCUSSION

All five operational performance indicators demonstrated significant improvements under AI-assisted operations compared to manual-only control. While the p-values (0.05) are at the conventional threshold of significance, the rank biserial correlation values ($r_s = 1.00$) across all variables show consistently strong and practically meaningful differences. These results provide robust support for the conclusion that AI-assisted operations enhance efficiency, consistency, and overall performance.

In the next phase of analysis, a Bayesian approach will be explored to further evaluate the strength of evidence, particularly in the context of small sample sizes. Subsequently, the Bayesian independent samples t-test was employed to assess group differences. For four parameters, the Bayes Factor (BF) exceeded 10, providing strong evidence of significant differences and confirming consistent directional trends across groups. In contrast, the parameter related to product yield yielded a Bayes Factor between 3 and 10, suggesting moderate evidence in support of the observed difference. Based on these statistical findings, the proposed hypothesis is considered to be supported.

The findings of this study provide several significant contributions to understanding the role of AI in industrial operations management. The comprehensive analysis across five critical performance dimensions reveals that AI assistance not only improves individual metrics but creates a synergistic effect that enhances overall operational excellence. The dramatic reduction in operational deviations (from 20-26 errors to 0-4 errors) represents more than statistical significance—it demonstrates AI's capacity to fundamentally transform operational discipline and safety culture.

From a practical standpoint, the results illuminate several key mechanisms through which AI enhances operator performance. First, the significant improvement in response time (Time to Target) suggests that AI support enables faster pattern recognition and decision-making during crisis scenarios. This is particularly critical in integrated systems where delayed responses can cascade into system-wide disruptions. Second, the enhanced product yield indicates that AI assistance helps operators optimize resource utilization even under stress, suggesting improved cognitive load management. Third, the superior utility efficiency demonstrates that AI support enables more sophisticated energy and resource management strategies that might be too complex for manual calculation during emergency situations.

The consistency of improvements across all measured dimensions is particularly noteworthy. The perfect effect sizes ($r_s = 1.00$) indicate that every operator performed better with AI assistance across every metric, suggesting that the benefits are not dependent on individual operator characteristics or specific scenario types. This universality of improvement provides strong evidence that AI assistance addresses fundamental limitations of human performance under stress rather than compensating for specific skill deficits.

However, several important considerations emerge from these findings. The detected versus undetected failure scenarios revealed interesting patterns in AI effectiveness. While AI assistance improved performance in both scenarios, the magnitude of improvement was greater for undetected failures, suggesting that AI's value is most pronounced when human operators have less preparation time and must rely more heavily on real-time analysis. This finding has important implications for AI system design, suggesting that emergency response capabilities should be prioritized in development efforts.

The Bayesian analysis provides additional insights into the strength of evidence supporting these conclusions. Bayes Factors exceeding 10 for four of five parameters indicate strong statistical support, while the moderate evidence for product yield (BF = 7.306) still represents meaningful improvement. The convergence of classical and Bayesian statistical approaches strengthens confidence in the practical significance of these findings despite the small sample size.

From a human factors perspective, the results challenge common concerns about AI reducing operator engagement or situational awareness. The consistent improvements across all metrics suggest that the AI system successfully augmented rather than replaced human decision-making capabilities. Operators remained actively engaged in the process while benefiting from enhanced analytical capabilities and decision support. This finding supports the development of collaborative AI systems rather than fully autonomous alternatives.

The implications for industry implementation are substantial. The quantified improvements provide a foundation for cost-benefit analysis of AI system deployment. For a typical integrated facility processing 400,000 barrels per day, the observed improvements in yield (15-20%) and utility efficiency (15-20%) could translate to millions of dollars in annual savings, easily justifying the investment in AI technology. Moreover, the dramatic reduction in operational errors has significant safety and environmental implications that extend beyond pure financial considerations.

This study provides a meaningful contribution to the field of industrial operations and human-machine collaboration by empirically demonstrating the performance impact of AI-assisted decision-making in real-time operational environments. The comparison between operator-only control and AI-assisted operations across five core performance metrics yields several noteworthy contributions:

1. Evidence-Based Validation of AI Benefits in Operations

While AI technologies are increasingly integrated into industrial systems, empirical evidence supporting their tangible benefits at the operator level is still limited. This research offers concrete, statistically supported evidence showing that AI assistance significantly improves key operational outcomes such as response time, yield, stability, utility usage, and inventory accuracy, even with minimal sample sizes. This validates the practical value of AI beyond theoretical or simulated environments.

2. Insight into Human-AI Synergy in Small-Scale Scenarios

Many studies on AI in process control focus on large datasets or complex algorithmic architectures. In contrast, this work emphasizes a small-sample, real-world scenario where human operators remain actively involved, and AI serves as a support tool rather than a replacement. This highlights the effectiveness of hybrid decision-making frameworks, offering valuable insights for industries where full automation is neither feasible nor desirable.

3. Methodological Rigor with Limited Data

The application of non-parametric statistics (Mann–Whitney U test) alongside effect size measures (rank biserial correlation) and Bayesian inference showcases a robust methodological framework suitable for studies with limited sam-ple sizes. This approach strengthens the reliability of conclusions drawn from small operational trials and offers a replicable model for other researchers dealing with similar constraints.

4. Performance Benchmarking for AI Adoption

The results provide a baseline performance benchmark for organizations considering AI integration into control rooms or production environments. By quantifying improvements across multiple dimensions with strong effect sizes, this study presents compelling justification for investing in AI tools that augment, rather than replace, human decision-making.

5. Contribution to the Human Factors and Technology Integration Literature

Finally, the study contributes to broader discussions on human factors engineering by showcasing how AI can enhance, not diminish, operator agency and effectiveness. The findings support a paradigm shift from automation versus human control to collaborative intelligence, where both entities work synergistically.

CONCLUSION

This study demonstrates that AI-assisted decision support significantly enhances operator performance in integrated refinery and petrochemical operations by improving response times, yields, error reduction, utility efficiency, and inventory management, even under high-stress shutdown scenarios. The consistency of improvements, confirmed by perfect effect sizes (rs=1.00rs=1.00) and strong Bayesian support (BF>10BF>10), indicates that AI addresses fundamental cognitive limitations rather than isolated operator deficiencies. These results provide robust quantitative justification for the adoption of human-AI collaboration in industrial operations, showing that such systems can deliver superior outcomes while preserving human oversight. Although limitations related to sample size and scenario diversity remain, the methodological rigor strengthens confidence in the findings. Future research should expand scenario variability, include larger and more diverse operator groups, and investigate long-term impacts of sustained AI use on human decision-making, adaptation, and trust in critical industrial environments.

REFERENCES

Bhattarai, K., & Yousef, M. (2025). Petroleum resources and energy transitions in the MENA region: Geopolitical and economic implications. In *The Middle East: Past, Present, and Future* (pp. 137–166). Springer.

Bigdeli, A., & Delshad, M. (2024). The evolving landscape of oil and gas chemicals: Convergence of artificial intelligence and chemical-enhanced oil recovery in the energy transition toward sustainable energy systems and net-zero emissions. *Journal of Data Science and Intelligent Systems, 2*(2), 65–78.

Brewer, B., Ashenbaum, B., & Blair, C. W. (2019). Cross-functional influence and the supplier selection decision in competitive environments: Who makes the call? *Journal of Business Logistics, 40*(2), 105–125. https://doi.org/10.1111/jbl.12212

Dozie, E. A., Johnson, N., Andrew, N. H., & Justine, A. C. (2023). Revolutionizing petrochemical production: Unleashing the full potential of Industry 4.0 to drive efficiency, harness reserve and propel innovation.

Hirsch, A., Parag, Y., & Guerrero, J. (2018). Microgrids: A review of technologies, key drivers, and outstanding issues. *Renewable and Sustainable Energy Reviews, 90*, 402–411. https://doi.org/10.1016/j.rser.2018.03.040

Jiang, L., Wu, Z., Xu, X., Zhan, Y., Jin, X., Wang, L., & Qiu, Y. (2021). Opportunities and challenges of artificial intelligence in the medical field: Current application, emerging problems, and problem-solving strategies. *Journal of International Medical Research, 49*(3), 03000605211000157. https://doi.org/10.1177/03000605211000157

Kaljunen, A. M. (2024). *Geopolitics and energy transition in the Baltic Sea region: A study on fossil road diesel market dynamics*.

Kerzner, H. (2025). *Project management: A systems approach to planning, scheduling, and controlling*. John Wiley & Sons.

Medina, C., Ríos, A. C. M., & González, G. (2022). Transmission grids to foster high penetration of large-scale variable renewable energy sources: A review of challenges, problems, and solutions. *International Journal of Renewable Energy Research (IJRER), 12*(1), 146–169.

Meegahapola, L., Mancarella, P., Flynn, D., & Moreno, R. (2021). Power system stability in the transition to a low carbon grid: A techno-economic perspective on challenges and opportunities. *Wiley Interdisciplinary Reviews: Energy and Environment, 10*(5), e399. https://doi.org/10.1002/wene.399

Ojuekaiye, O. S. (2024). Petroleum industry value chain optimization: The inevitability of midstream and downstream development. In *Asset Management and Information* (Paper D031S022R007). SPE Nigeria Annual International Conference and Exhibition.

Rao, Y. S. (2023). *Production and operations management*. Academic Guru Publishing House.

Rojas, L., Peña, Á., & Garcia, J. (2025). AI-driven predictive maintenance in mining: A systematic literature review on fault detection, digital twins, and intelligent asset management. *Applied Sciences, 15*(6), 3337. https://doi.org/10.3390/app15063337

Sharma, K. L. S. (2016). *Overview of industrial process automation*. Elsevier.

Swink, M., & Schoenherr, T. (2015). The effects of cross-functional integration on profitability, process efficiency, and asset productivity. *Journal of Business Logistics, 36*(1), 69–87. https://doi.org/10.1111/jbl.12079

Tien, J. M. (2017). Internet of things, real-time decision making, and artificial intelligence. *Annals of Data Science, 4*(2), 149–178. https://doi.org/10.1007/s40745-017-0110-8

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., & Qiu, C. W. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation, 2*(4), 100179. https://doi.org/10.1016/j.xinn.2021.100179