The Benchmarking Trap

How composite cost targets can destabilise underground mines

From peer comparison to balanced capacity, chronic firefighting and higher unit cost

Hendrik Lourens, Stratflow Australia, FAusIMM | September 2026

Executive Summary

Benchmarking can be useful when it is used as a structured inquiry: compare like with like, understand why performance differs, adapt relevant practices and test whether they improve value. The failure examined in this paper begins when benchmarking stops being inquiry and becomes a resource-setting mechanism.

The critical failure mode is the construction of a synthetic "best mine" that has never existed. The lowest development cost may come from Mine A, the lowest stoping cost from Mine B, the lowest materials cost from Mine C and the smallest maintenance complement from Mine D. When these separate minima are assembled into one corporate standard, the resulting target does not represent any operating system that has demonstrated joint feasibility.

If each department is then required to close its apparent cost gap - often by reducing people, equipment, spares, inventory or standby capacity - while the mine remains committed to a production plan derived from average rates, the intervention pushes every link toward "just enough" capacity. Arrie van Niekerk described the result as a balanced capacity chain: every activity appears efficient in isolation, but the chain has too little recovery capacity to absorb normal variation. Delays propagate, the system constraint is starved or blocked, bottlenecks appear to move, output becomes volatile and unit cost rises as fixed costs are spread over fewer saleable tonnes.

Underground mines are especially exposed because the cost and productive capacity of an activity depend on mining method, depth, travel distance, ground conditions, geometry, grade, sequencing, ventilation, materials logistics and access to work fronts. Published mining studies confirm that these contextual drivers materially change cost and productivity, even among mines using the same broad method. Operations research also confirms that serial systems with limited buffers are governed by starvation, blockage and interdependence rather than by the arithmetic average of each part.

Central finding: The problem is not comparison. It is converting decontextualised, function-level minima into simultaneous cost and resourcing targets without proving that the combined design can deliver mine-wide flow under variability.

The literature does not yet directly test the complete causal proposition advanced here. The evidence is strongest for its component links; the full mining mechanism remains a testable hypothesis supported by systems logic, operational mathematics, Arrie van Niekerk's teaching, documented Stratflow experience and a consistent body of practitioner observations. This paper therefore separates established evidence from interpretation and proposes practical safeguards and a research agenda.

1. What benchmarking is meant to do

Professional benchmarking guidance treats the exercise as a learning process, not a ranking contest. APQC describes a sequence of planning, collection, analysis and adaptation: performance data locate a gap, while practice research is used to understand the drivers and enablers behind the result. Findings create value only when they are adapted to the receiving organisation [1].

Mining-specific guidance makes the same point. Hall and Harper describe benchmarking as a way to identify performance gaps and prioritise improvement, but emphasise that it must be tied to strategic goals and measurable value [2]. This distinction matters. A benchmark can be a question - "what explains the difference?" - or an instruction - "remove the difference." Only the first preserves the possibility that the observed gap is caused by geology, mine design, production scale, work mix, risk exposure or a deliberate investment in system resilience.

Benchmark as inquiry	Benchmark as target	Operational consequence
Locates an unusual difference	Defines the difference as waste	The explanation is decided before technical investigation.
Studies the practices and conditions behind performance	Copies the observed number	Context and trade-offs disappear.
Adapts a practice to the receiving mine	Standardises the input or complement	The mine is forced toward a generic design.
Tests the effect on whole-mine value	Tracks departmental variance to target	A local saving can be booked while mine throughput falls.

Table 1. The same comparison produces very different behaviour depending on how it is used.

2. Why underground mines are unusually difficult to benchmark

The common denominator "underground mine" conceals systems that are not operationally equivalent. Open stoping, narrow-vein, cut-and-fill, room-and-pillar, block caving, sublevel caving and longwall mining do not use the same production logic. Even mines using the same method differ materially in orebody geometry, rock mass behaviour, depth, access, haul distance, ventilation, seismic or geotechnical controls, development intensity, backfill requirements, dilution, recovery and available work fronts.

The published evidence is consistent with the practitioners in Joel Gray's discussion. Ulrich, Trench and Hagemann analysed 57 Australian gold mines and found clear grade and scale effects, with grade-cost relationships that differed radically between mines; they also identified geometry and mineralisation style as relevant cost drivers [3]. A global review of 21 sublevel-caving mines found significant differences in practices and productivity, and explicitly analysed the effects of mine scale, depth and ground conditions [4].

Recent underground mine-design research adds a sharper warning about average unit costs. Valencia, Morales Varela and Gamache show that a constant cost per tonne cannot adequately represent shared infrastructure and transport distance; the gap between assumed input cost and designed output cost can be substantial [5]. If a constant cost model can distort optimisation within one mine, an unadjusted cost from another mine is an even weaker design instruction.

Conditions that can make an apparently high cost rational

- Long travel and materials routes that reduce effective face time.
- Difficult ground, support or rehabilitation requirements that protect access and people.
- A higher development burden required to preserve future production flexibility.
- Additional crews, mobile equipment, spares or inventory held to recover from disruption.
- A different constraint location, making high utilisation valuable in one area and harmful in another.
- Grade, dilution, recovery or haulage differences that change the denominator of the cost metric.
- More demanding ventilation, refrigeration, pumping, backfill or seismic controls at depth.

Underground reality: Two mines can report the same activity name while performing technically different work under different constraints. A comparable label is not evidence of a comparable operating system.

3. The mine that never existed

Arrie van Niekerk illustrated the problem with five underground mines. Suppose Mine 1 reports the lowest development cost, Mine 3 the lowest stoping cost and Mine 4 the lowest cleaning cost. The lowest figure in each row becomes the benchmark for every other mine. The composite looks rigorous because every number has been achieved somewhere. Yet no mine has achieved all of the numbers together.

Function	Selected minimum	What disappears	What the standard implies
Development	Mine 1	Distance, ground, headings, future optionality	Match Mine 1 input and unit cost

Function	Selected minimum	What disappears	What the standard implies
Stoping	Mine 3	Method, stope geometry, fill, dilution, sequencing	Match Mine 3 input and unit cost
Cleaning / haulage	Mine 4	Tramming distance, congestion, shaft interface	Match Mine 4 input and unit cost
Maintenance	Mine 2	Fleet age, criticality, standby philosophy, spares	Match Mine 2 complement and inventory
Services / support	Mine 5	Depth, pumping, ventilation, risk controls	Match Mine 5 support cost

Table 2. A synthetic benchmark can be assembled from real numbers without representing a jointly feasible mine.

This is a form of selection bias at system level. Each minimum is conditioned by the rest of the mine that produced it. The low stoping cost may depend on generous development capacity, an unusually favourable work mix or a maintenance resource that appears expensive elsewhere. Separating the number from those enabling conditions converts an outcome into an input standard.

Research on formal benchmarking models reinforces the need for caution. Guevel, Ramon and Aparicio argue that benchmarks must account for the circumstances of the unit being evaluated and that targets must be achievable; even mathematically "closest" targets can remain impractical when they require extreme movement in a variable [7]. The composite-minimum approach is less conservative: it can require extreme movement in every function at once.

4. How the benchmark becomes a balanced capacity chain

The causal mechanism is not the benchmark spreadsheet by itself. It is the sequence of management decisions the spreadsheet legitimises. The chain below synthesises Arrie's lecture, the ERP and budgeting paper by Urban and Lourens, the attached Stratflow material and the practitioner discussion [10-12].

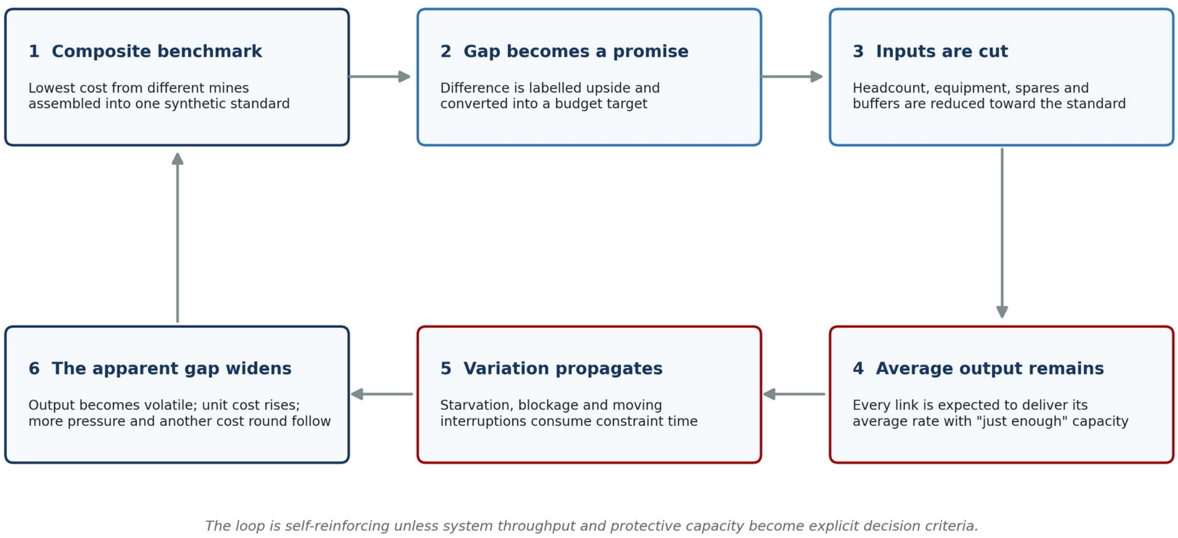

Figure 1. The self-reinforcing path from a composite benchmark to a balanced capacity chain.

Step 1: Separate minima are assembled

The best observed cost for each activity is selected from different mines or periods. Correlations and trade-offs between the activities are not preserved. The target describes a portfolio of parts, not a demonstrated production system.

Step 2: The gap is monetised

The difference between actual performance and the composite is labelled an opportunity. Because the gap is visible and quantifiable, it is attractive to boards and transformation offices. The uncertainty surrounding causation is compressed into a precise savings promise.

Step 3: The number is converted into inputs

To close the gap, functions reduce the resources that appear to explain it: people, contractor hours, mobile equipment, maintenance coverage, spares, inventory, work fronts or buffer stocks. Each reduction is rational inside the local cost centre.

Step 4: The production promise remains

The mine is still expected to deliver a plan calculated from average rates. Budgeting works backwards from the desired saleable tonnes and asks how much of each resource is required, on average. After benchmarking, the answer tends toward the smallest apparently adequate complement - Arrie's "just enough of everything" chain.

Step 5: Protective capacity disappears

Non-constraint resources lose the capacity to recover after disruption. Buffers become suspect because they look like inventory, underutilised equipment or excess labour. The resource that is idle today is removed even if that idle time is what allows it to catch up tomorrow.

Step 6: Variation and interdependence take control

A delay in development removes a future work front; a materials delay prevents a blast; a maintenance delay idles the loading or hoisting constraint; an unavailable stope changes the blend; congestion blocks downstream movement. Without decoupling buffers, each loss is passed on while local gains cannot be stored. The mine experiences moving interruptions rather than one stable, manageable constraint.

Step 7: The response repeats the cause

Output falls below plan and cost per tonne rises. The underperformance is read as proof that the original gap remains. More dashboards, pressure, action lists and cost reduction follow. The intervention becomes self-validating because its own consequences are interpreted as evidence that the mine has not yet complied strongly enough.

5. The operating physics: why averages do not add up

A serial mine-production chain is governed by the interaction between capacity, variability, buffers and dependency. A budget that assigns every step an average capacity of 10 units per hour does not create a reliable 10-unit-per-hour system. When any necessary link is unavailable, work can starve downstream activities or block upstream activities. Limited buffers reduce the ability of one part of the chain to continue while another is disturbed.

Modern production-line research models these effects explicitly. Calvo's analysis of heterogeneous serial lines with limited buffers describes how blocking and starvation reduce effective process rates and how the effect can propagate recursively through the chain. The same paper notes the "bowl phenomenon": under some conditions, an intentionally unbalanced work allocation produces more throughput than a perfectly balanced distribution with the same work content [6]. That is directly relevant to the belief that equalised capacity is automatically optimal.

Independent serial links	Availability of each link	Illustrative simultaneous availability	Maximum rate in a 10-unit/h chain
1	90%	90.0%	9.00 units/h
3	90%	72.9%	7.29 units/h
6	90%	53.1%	5.31 units/h
6	80%	26.2%	2.62 units/h

Table 3. A deliberately simplified no-buffer illustration: p^n. It demonstrates compounding, not a forecast for a real mine.

Important mathematical caution: The p^n calculation assumes independent links that must be simultaneously available, no effective buffers, no parallel capacity and a common nominal rate. Real mines require discrete-event or other system modelling. The illustration is retained because it makes the direction of the error visible; it must not be presented as an empirical production estimate.

The more useful operational measures are therefore not local utilisation alone but constraint run time, starvation and blockage minutes, buffer health, schedule adherence, recovery time after disruption, cumulative interdependent delay and the distribution of daily saleable tonnes. These measures reveal whether spare capacity is waste or insurance.

6. What miners report experiencing

Joel Gray's article and the attached discussion provide a concentrated practitioner account of what benchmarking programmes feel like underground. These comments are not a representative sample and do not establish causation. They are nevertheless useful qualitative evidence because the same themes recur across technical, operational, board and assurance roles.

False comparability. Mines sharing an industry label are treated as equivalent despite different methods, geology, depth, conditions and plans.

Technical capability gap. Review teams are described as dominated by accountants and data specialists with insufficient underground mining or mine-planning depth.

Leadership distance. Boards and executives without relevant operating exposure can prefer a precise benchmark narrative to a technically qualified explanation.

Pressure transferred to site. The numerical gap becomes a demand for more effort from a workforce already perceived as under-resourced.

Irrelevant KPIs and standardisation. Measures from bulk, coal or surface contexts are imported into technically different underground methods.

Low trust and alienation. Repeated programmes are reported to cost heavily, deliver little and leave site teams less willing to engage with the next initiative.

Evidence asymmetry. Frontline technical practitioners report a persistent problem but struggle to obtain data in a form that satisfies the corporate demand for proof.

Target-setting research provides one adjacent empirical warning. Bouwens and Kroos found that retail managers reduced sales activity when favourable current performance was likely to ratchet up future targets [8]. The setting is not mining, but the behavioural mechanism matters: when exceptional performance becomes the next compulsory baseline, people may protect themselves by withholding, re-timing or gaming performance. A record or peer result is not a neutral target.

What the comments add: They do not prove the balanced-capacity mechanism. They show that experienced practitioners recognise its organisational symptoms: irrelevant comparisons, technical knowledge displaced by numerical authority, pressure on site, weak learning and repeated loss of trust.

7. Consequences for the mine

Lower throughput and higher cost per tonne

Many underground mines have high fixed costs. When output falls, the apparent saving in a department can be overwhelmed by the lost contribution of saleable tonnes and the spreading of fixed cost over a smaller denominator. This is why removing a maintenance resource, standby machine or materials buffer can lower the cost-centre budget while increasing mine-wide unit cost.

Moving bottlenecks and permanent firefighting

When several links have similar effective capacity and little ability to recover, the most visible problem changes from shift to shift. Management follows yesterday's red number, action lists roll forward and technical attention is consumed by escalation. The system has many interruptions but no stable, deliberately managed constraint.

Plans that become obsolete early

A plan built on average activity rates without sufficient work fronts, time, material or equipment buffers can be internally coherent and operationally fragile. One missed dependency forces resequencing, which changes the conditions for other departments and creates further variance. The plan is repeatedly repaired rather than executed.

Departmental success and mine-wide failure

Local cost and utilisation measures can remain green while the mine misses saleable tonnes. Each department has evidence that it complied; the loss exists in the interactions. Because no cost centre owns starvation, blockage or cumulative interdependent delay, the largest system loss can remain unpriced.

Erosion of technical judgement

A benchmark turns an uncertain technical judgement into an apparently objective variance. Once the number acquires authority, explaining why the receiving mine is different can sound defensive. Over time, people learn that the safest response is to comply with the template, document exceptions and protect their department rather than challenge the system design.

Workforce pressure, morale and safety precursors

Under-resourcing and repeated resequencing can increase reactive work, handover complexity, overtime, fatigue and unplanned interactions. The JINSE perspective by Lourens argues that operating instability may reduce the time and organisational capacity available for maintenance, critical-control verification and response to weak signals [9]. That relationship is plausible but not established by the benchmarking evidence reviewed here; it should be treated as a priority research question, not a proven safety effect.

8. Evidence map: what is established and what is proposed

Proposition	Current support	Evidence status
Underground costs are context-dependent.	Mining studies identify grade, scale, geometry, mineralisation style, depth, ground conditions, infrastructure and transport distance as material drivers [3-5].	Strong support
Good benchmarking requires driver analysis, adaptation and achievable targets.	Professional guidance and benchmarking research explicitly require comparable definitions, contextual understanding and practical attainability [1,2,7].	Strong support
Limited buffers and unreliable serial links create starvation, blockage and throughput loss.	Operations-research literature models these mechanisms directly [6].	Strong support
Targets ratcheted from observed performance change behaviour.	Empirical evidence exists in other industries [8].	Adjacent support
Mining benchmarking programmes often alienate site teams and import irrelevant KPIs.	Consistent practitioner observations in the Gray discussion; not a representative study [10].	Qualitative support
Composite functional minima plus resource cuts cause balanced capacity in underground mines.	Mechanism described by van Niekerk and Urban/Lourens, consistent with systems theory and practitioner accounts [11,12].	Novel, testable hypothesis
The resulting instability degrades safety-control execution.	Proposed in the JINSE perspective; direct empirical test not yet available [9].	Research hypothesis

Table 4. Evidence boundary for the paper's central argument.

9. A safer way to use benchmarks

Benchmarking should not be abandoned. It should be prevented from silently becoming an operating-system design. The following safeguards preserve its learning value while protecting mine-wide flow.

Define the decision first. State whether the benchmark is for learning, target setting, resourcing, capital allocation or assurance. A comparison made for learning should not automatically authorise a cost cut.

Establish comparability. Segment by mining method, depth, geometry, ground conditions, production scale, work mix, development intensity, haul distance, grade and relevant risk controls. Record what remains non-comparable.

Preserve joint feasibility. Do not assemble function-level minima into a simultaneous target unless one operation has demonstrated the combination or a validated system model shows it is feasible.

Trace the enabling conditions. For every low-cost result, identify the work fronts, buffers, standby capacity, maintenance support, inventory, capital intensity and upstream investments that enabled it.

Model the whole chain. Test proposed resource changes against saleable throughput, variability, constraint starvation/blockage, recovery time and cost per tonne - not only against the local budget.

Identify the system constraint. Specify which activity sets mine-wide throughput. High utilisation is required there; elsewhere utilisation is subordinate to protecting flow.

Name and price protective capacity. Document why non-constraint crews, equipment, spares, space or inventory exist and value them against the cost of lost constraint time.

Use ranges and scenarios. Replace single-point certainty with credible ranges for geology, availability, cycle time and schedule performance. State the probability or operating conditions attached to the target.

Require technical sign-off. Frontline technical services, operations, maintenance, finance and safety should jointly test whether the proposed target is physically and organisationally coherent.

Stage the change. Pilot changes at limited scope. Stop or redesign the intervention if throughput distribution worsens, variability rises, maintenance backlog grows or recovery capacity is consumed.

10. Questions for boards, owners and mine executives

1. Which real mine has demonstrated the complete benchmark combination, not merely each component separately?
2. What technical and operating conditions explain the selected comparator's result, and which of those conditions are absent here?
3. Which resources are being removed, and what disruption or recovery function does each currently perform?
4. Where is the stable system constraint, and how will the proposal affect its starvation, blockage and run time?
5. What does the proposal do to the distribution of daily saleable tonnes, not only to average cost?
6. What mine-wide value is lost if one hour of constraint time is sacrificed to achieve the departmental saving?
7. What evidence would cause management to stop, reverse or redesign the programme?
8. Have frontline technical services been asked to explain the gap before the savings target was announced?

The governing test: A benchmark recommendation should be rejected as incomplete if it cannot explain how the proposed cost and resource configuration will deliver stable mine-wide flow under normal variability.

11. A research agenda

The central proposition can be tested using before-and-after data from benchmarking or "full potential" programmes across a portfolio of underground mines. The independent variable would be the degree to which departmental resource levels are moved toward external or internal minimum-cost standards without explicit allowance for protective capacity. The outcomes should include more than cost variance.

- Daily saleable tonnes, coefficient of variation, P10-P90 range and plan attainment.
- Constraint starvation and blockage minutes; buffer penetration and replenishment time.
- Number and duration of plan changes; schedule churn and reactive supervisory work.
- Maintenance backlog, stock-outs, deferred work, equipment standby coverage and overtime.
- Cost per tonne, incremental contribution, EBITDA and working capital.

- Leading safety measures such as delayed verification, precursor events and exposure created by unplanned work.

A credible design would compare mines that received a composite, cost-led benchmark intervention with matched mines that used context-adjusted, system-level benchmarking. Interrupted time-series or difference-in-differences methods could test whether changes in resource intensity preceded changes in flow stability. Qualitative interviews should be used to identify the decision pathway, not as a substitute for operating data.

A null or boundary-limited result would be valuable. The hypothesis may apply strongly to tightly coupled underground methods and weakly to operations with large natural buffers, excess work fronts or low fixed-cost exposure. The aim should be to discover where the mechanism operates, not to validate a predetermined conclusion.

Conclusion

Benchmarking is most dangerous when it looks most precise. A low cost observed in one department at one mine is real, but the conclusion drawn from it may not be. The number is the outcome of a particular orebody, mine design, constraint, work mix, risk environment and set of supporting resources. Removing that context and combining the number with minima from other mines creates a synthetic operation that has never run a shift.

When the synthetic standard is converted into simultaneous resource cuts while the production plan continues to assume average capacity, the mine is pushed toward a balanced capacity chain. Protective capacity is removed because it looks idle, buffers are reduced because they look costly, variation propagates and the constraint loses time. The local saving is visible immediately; the lost throughput arrives later and is attributed elsewhere.

The appropriate response is not to reject benchmarking but to restore its proper role. Use it to identify questions, understand enabling practices and challenge assumptions. Do not allow it to set the architecture of an underground mine without a system model, technical explanation and evidence that the combined target is feasible under variability.

Final message: Benchmark systems, not isolated parts. Compare to learn. Resource to protect flow. Judge the result at the point of sale.

Appendix A. Complete note of the attached LinkedIn comments

The following register records every comment and author reply contained in the supplied document "Joel Gray evil of benchmarking and comment.docx". The wording is condensed to preserve the substantive point. Reactions such as "low hanging fruit" and "Hardrock Socialism" are included because the instruction was to note all comments, but they are treated as sentiment rather than technical evidence.

Contributor	Substantive note	Relevance to this paper
Hendrik Lourens	Benchmarking imposes certainty on an uncertain system, treats different mines as comparable and turns the apparent gap into pressure on an under-resourced workforce.	Supports context, systems redesign and plan improvement rather than more effort.
Sam Patterson	Compared benchmarking unlike underground mines to comparing Aston Martin and Nissan because both make cars. Suggested many boards and executives lack mining-operations depth.	Comparability and governance capability are central risks.
Joel Gray - reply to Patterson	Observed bulk-mining leaders importing irrelevant coal-style KPIs. Said efficiency is a large lever in bulk methods but the opportunity in open stoping, narrow vein and cut-and-fill often lies elsewhere.	Mining method changes the leverage point; one KPI architecture should not span all methods.
Bryan Pullman	Reported BI teams with accountants, data scientists and engineers lacking mining expertise; questioned how teams without relevant experience solve deep technical problems.	Benchmark analysis requires underground technical capability, not data capability alone.
Keith Whitchurch	Called internal benchmarking particularly flawed and recommended Stephen Shapiro's Best Practices Are Stupid.	Internal peers can still be incomparable; "best practice" needs explanation and adaptation.
Alex Atkins	Linked the issue to her PhD interviews and said frontline technical practitioners struggle to persuade leaders without formal evidence.	Identifies the evidence gap and the need for operational data.
Alex Doll	Used humour to list non-standard geology, orebody shape, depth, grade, rock behaviour, grindability, recovery and dewatering as reasons standard BI comparisons fail.	Provides a concise catalogue of contextual variables erased by standardisation.
Joel Gray - reply to Doll	Humorously agreed with the geological critique.	Affirms the context argument; no additional evidence.
Alan Watt	Said the problem extends beyond mining; only consultants who first understood the company and operation delivered value.	Sequence matters: understand operations before analysing gaps.
Joel Gray - reply to Watt	Acknowledged the cross-industry experience while limiting his own claim to mining.	Appropriately bounds the original article's claim.
Richard Thomas	Described turnaround work as "churn things around."	Signals cynicism about repeated transformation cycles; sentiment, not causal evidence.
Lionel Kelly	Reported enduring many BI strategies over 44 years; said they cost millions, achieved little and alienated people on site.	Supports the recurring themes of low sustainability, high disruption cost and loss of trust.
Coogee Barbuzza	Characterised benchmarking as "low hanging fruit."	Suggests the approach is selected because the analysis is easy to sell, not because implementation is most valuable.
William Jud	Labelled the practice "Hardrock Socialism."	Expresses opposition to imposed uniformity; rhetorical sentiment rather than technical evidence.

Table A1. Complete register of the comments in the supplied LinkedIn discussion.

Appendix B. Source and method note

This white paper is a synthesis, not a systematic review. The web search prioritised mining conference papers, peer-reviewed research, professional benchmarking guidance and primary publication pages. The supplied documents were treated as practitioner and internal sources. Claims were classified by evidence strength, and the paper does not present the complete benchmarking-to-balanced-capacity pathway as empirically proven.

The simplified p^n example was recalculated for mathematical accuracy and explicitly bounded by its assumptions. It should not be substituted for discrete-event simulation, mine scheduling analysis or a site-specific flow model.

References

1. APQC (2026). What does it mean to benchmark a process or performance?
2. Hall, A. J. & Harper, P. J. (2005). Benchmarking - A Practical Technique for Measuring and Improving Operational Performance. Ninth Underground Operators' Conference, AusIMM.
3. Ulrich, S., Trench, A. & Hagemann, S. (2017). Towards an understanding of production costs at Australia's underground gold mines. 13th AusIMM Underground Operators' Conference.
4. Campbell, A. D. (2023). A global review of productivity and operational performance in sublevel caving mines. Underground Operators Conference, AusIMM.
5. Valencia, S., Morales Varela, N. & Gamache, M. (2026). Maximizing Underground Mine Reserves Through a More Accurate Cost Model and the Application of Machine Learning. Natural Resources Research, 35, 2335-2351.
6. Calvo, R. (2026). Identification of bottlenecks and optimization of throughput for manufacturing serial lines with limited buffers. International Journal of Advanced Manufacturing Technology.
7. Guevel, H. P., Ramon, N. & Aparicio, J. (2025). Benchmarking in data envelopment analysis: balanced efforts to achieve realistic targets. Annals of Operations Research, 351, 1403-1426.
8. Bouwens, J. & Kroos, P. (2011). Target ratcheting and effort reduction. Journal of Accounting and Economics, 51(1-2), 171-185.
9. Lourens, H. (2026). Breaching the Mining Fatalities Plateau: Financial Pressure, Operational Instability, and the Case for Flow-Based Safety. Journal of Industrial Safety.
10. Gray, J. (2026). Why Benchmarking Fails Underground Mining. LinkedIn article and attached discussion, 31 August 2026. Supplied source document.
11. van Niekerk, A. (undated). Thabazimbi TOC Flow lecture transcript. Supplied source document.
12. Urban, B. & Lourens, H. (2016). Dealing with the Worldwide Decline in Mining Productivity: A focus on the underlying paradigm and the ERP Budgeting Systems maintaining the decline. Supplied source document.
13. Stratflow Australia (2026). The Best-Practice Trap; Predictability Premium series; and The Dual Operating System. Supplied source documents.