Volume 01 / Soft Sensor Architectures for Mineral Processing Control

A two-paper compendium on physics-informed soft sensors that close the latency between laboratory measurement and real-time process control. Designed as feedforward inputs to the Axiom hybrid APC layer, with operator-approved supervisory optimisation above.

AUTHORS	AFFILIATION	CLASSIFICATION
T. Davhana	**Axiom Process Technology**	**Public Release**
D. Mthimunye	**Nelspruit, South Africa**	**Volume 01**

CONTENTS

P-001 A Physics-Informed Soft Sensor for Real-Time Ore Hardness Estimation in SAG Mills

P-002 An Ensemble Soft Sensor for SiO_2 Grade Prediction in Iron Ore Reverse Cationic Flotation

FOREWORD

These two papers establish the layer of the Axiom control stack that converts laboratory latency into a continuously available, physically grounded signal. P-001 addresses the comminution circuit; P-002 addresses the flotation circuit. Both rest on the same discipline: combine first-principles physics with machine-learning corrections, validate against the plant's own laboratory turnaround, and feed the result forward into the Advanced Process Control layer rather than displaying it on a dashboard.

The two papers are produced by Axiom for public release and may be redistributed in unmodified form. The underlying methods are deployed under engagement on industrial sites; the papers describe the architecture, not the proprietary tuning that the engagement delivers.

A Physics-Informed Soft Sensor for Real-Time Ore Hardness Estimation in SAG Mills

A residual-based architecture combining the Morrell mill power model with Extra Trees regression, delivering a continuous, SCADA-only hardness signal suitable for MPC feedforward.

AUTHORS	AFFILIATION	CLASSIFICATION
T. Davhana	**Axiom Process Technology**	**Public Release**
D. Mthimunye	**Nelspruit, South Africa**	**Vol 01, No 01**

ABSTRACT

Ore hardness drives SAG mill power draw, throughput, and downstream particle size distribution, yet plants typically infer hardness only through periodic Bond Work Index or SPI laboratory tests at intervals of hours to days. The latency between hardness change and operator response is the single largest source of avoidable variability in comminution. This paper sets out a two-method architecture for real-time hardness estimation from existing SCADA infrastructure, without requiring additional instrumentation. Method 1, the SCADA-only reference implementation, uses the Morrell mill power model to compute the power the mill *should* draw and treats the residual against actual power as the hardness signal. An Extra Trees regressor is trained on this residual together with engineered features and periodic Bond Wi or SPI labels, yielding a continuous hardness estimate that updates every DCS scan cycle. Method 2 extends the architecture to a geospatial mine-to-mill digital twin, fusing blast block models, shovel GPS, and stockpile blending into a thirty-minute look-ahead forecast that serves as the prior in a Kalman filter, with the physics residual as the measurement update. The architecture is sensor-light, physically grounded, and directly consumable by Model Predictive Controllers as a feedforward signal.

KEYWORDS

Soft sensor, ore hardness, SAG mill, Morrell model, Extra Trees regression, Kalman filter, mine-to-mill, geospatial digital twin, MPC feedforward.

1. Introduction

The economic performance of a mineral processing plant is dominated by the comminution circuit, both in energy intensity and in its determination of the particle size distribution presented to flotation. Among the variables that drive comminution behaviour, ore hardness is simultaneously the most consequential and the least directly observable. Plants commonly characterise hardness through SAG Power Index or Bond Work Index assays, but the laboratory turnaround between sample and result, measured in hours to days, makes these labels operationally retrospective. By the time hardness is known, the ore has moved.

The cost of this latency is variability. When a hardness transition arrives at the mill unannounced, operators discover it through reduced throughput, increased mill load, or departures in cyclone overflow particle size. The corrective adjustments to feed rate, mill speed, or water addition come after the disturbance has propagated through the circuit. Where supervisory or Model Predictive Control is in place, the controller is forced to react to its own state estimates rather than to a leading indicator of the incoming ore. Variance reduction at the mill therefore requires that hardness become a measured, or at least estimated, real-time variable.

This paper presents a two-method soft sensor architecture for real-time hardness estimation. The architecture is designed around two non-negotiable constraints. First, no additional instrumentation is required beyond what a typical SAG mill installation already provides through its Distributed Control System. Second, the output must be a continuously updated, physically interpretable signal of a form that can be consumed directly as a feedforward input by an MPC or Advanced Process Control layer. The architecture is the SCADA-only reference (Method 1) followed by a geospatial extension that adds look-ahead (Method 2).

2. Method 1, the SCADA-Only Reference Implementation

2.1 Physics layer, the Morrell mill power model

The Morrell (1996) C-model expresses the gross power drawn by a tumbling mill as a function of geometric and operational variables that are independent of ore hardness:

$$P_{calc} = K \cdot D^{2.5} \cdot L \cdot \rho_c \cdot J_c \cdot \phi \cdot f(\phi)$$

where D and L are mill diameter and length, ρ_c is the charge density, J_c is the fractional charge filling inferred from bearing pressure, ϕ is the fractional speed expressed as a fraction of critical, and K is a calibration constant capturing liner and lifter geometry. The model is well established in the comminution literature and is the standard against which mill power is reconciled in plant design and audit.

The interpretive move is the residual. Once geometry, charge, and speed are accounted for, the difference between the actual power drawn and the model prediction is dominated by the energy required to break the ore. Harder ore breaks at higher specific energy; softer ore at lower. The residual therefore serves as a continuously available, physically grounded hardness signal:

$$P_{residual} = P_{actual} - P_{calc} \rightarrow \text{ore hardness contribution}$$

2.2 Feature engineering and the prediction model

The hardness residual alone is informative but noisy. The full feature vector presented to the regression layer combines the residual with engineered process features that capture the temporal dynamics of the mill. The hardness residual P_actual – P_calc is the principal feature. A normalised load term, computed as bearing-pressure-derived load divided by the square of mill speed, decouples mass from velocity and stabilises the residual against speed transitions. Specific-energy lag windows at five, ten, and thirty minutes capture the hardness trend over operationally relevant timescales.

The prediction model is an Extra Trees regressor (Geurts, Ernst & Wehenkel, 2006), chosen for its noise robustness, its insensitivity to feature scaling, and its ability to handle the skewed residual distributions characteristic of mill power data. The model is trained against the periodic SPI or Bond Wi laboratory labels available from the plant. The training target is a calibrated hardness index in the units of the available label, with held-out validation requiring an R^2 of at least 0.85 before the sensor is signed off for control use.

2.3 Output and integration with control

The output of Method 1 is a continuous, scalar estimate of ore hardness that updates every DCS scan cycle, typically one second. The signal is presented to the MPC layer as a measured disturbance, allowing the controller to adjust feed rate, mill speed, and water addition in anticipation of the hardness change rather than in reaction to its downstream consequences. The sensor is sensor-light by design: it requires only signals already present on every modern SAG mill installation.

3. Method 2, the Geospatial Mine-to-Mill Digital Twin

3.1 Motivation, the look-ahead horizon

Method 1 estimates hardness from the present state of the mill. It cannot estimate hardness that has not yet arrived at the mill. For circuits where the residence time from blast to mill is long enough that a look-ahead horizon adds value, typically thirty minutes to several hours depending on stockpile management, the architecture extends naturally to a mine-to-mill digital twin.

3.2 Geospatial fusion layer

The fusion layer combines three upstream data streams. The blast block model provides lithology and Rock Quality Designation by block identifier from pre-classified drilling. Shovel GPS, typically at RTK accuracy of ten to thirty centimetres, gives the real-time location of mass extraction. The geometallurgical database supplies SPI, Bond Wi, and ore density per block. For each conveyor segment between the loading point and the mill, the residence-time model computes the temporal lag $\Delta t = L \div v$, and a stockpile blending model tracks the resulting multi-shovel mass packets first-in-first-out through the stockpile. The output is a forecast hardness vector for the ore that will arrive at the mill in the next T+n minutes.

3.3 Kalman filter and hybrid state estimate

The geospatial forecast is the prior in a Kalman filter; the physics residual from Method 1 is the measurement update. The Kalman gain K weights geological uncertainty against real-time SCADA evidence at every scan cycle. Where the geospatial signal is confident, for example shortly after a well-characterised blast, the forecast dominates and the controller receives a clean look-ahead hardness signal. Where the geospatial signal is uncertain, for example through complex stockpile reclaim, the measurement residual dominates and the architecture degrades gracefully to Method 1

behaviour. The controller never loses its hardness input.

4. Architectural Position in the Axiom Control Stack

The hardness soft sensor sits in the third layer of the Axiom control hierarchy: above the healthy DCS, above the anomaly detection layer that flags sensor dropout and sticky-valve behaviour, alongside the other physics-informed soft sensors that generate feedforward signals for the Advanced Process Control layer. The APC layer, a hybrid of Fuzzy Logic Control for constraint management and Model Predictive Control for multivariable coordination, consumes the hardness signal as a measured disturbance. The AI Optimisation Layer above the APC computes the economic setpoints into which the hardness feedforward is embedded. The contribution of the soft sensor is therefore leverage on every layer above it, not a standalone output.

5. Conclusion

Real-time ore hardness estimation is achievable from existing SCADA infrastructure without additional instrumentation. The architecture presented in this paper combines the Morrell mill power model with Extra Trees regression to convert the unexplained portion of mill power draw into a continuous, physically grounded hardness signal. Where the upstream mine and stockpile management are sufficiently instrumented, the architecture extends naturally to a thirty-minute look-ahead forecast through Kalman fusion of geospatial and physics signals. In both forms, the output is a feedforward signal directly consumable by the APC layer above it. The cost of hardness latency becomes addressable.

References

Geurts, P., Ernst, D. and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), pp. 3–42.

Morrell, S. (1996). Power draw of wet tumbling mills and its relationship to charge dynamics. *Transactions of the Institution of Mining and Metallurgy, Section C*, 105, pp. C43–C53.

Napier-Munn, T.J., Morrell, S., Morrison, R.D. and Kojovic, T. (1996). *Mineral Comminution Circuits: Their Operation and Optimisation*. JKMRC Monograph Series in Mining and Mineral Processing 2. JKMRC, Brisbane.

Wills, B.A. and Finch, J.A. (2016). *Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery*, 8th ed. Butterworth-Heinemann, Oxford.

An Ensemble Soft Sensor for SiO_2 Grade Prediction in Iron Ore Reverse Cationic Flotation

A T+30 minute concentrate-grade prediction architecture stacking LSTM, 1D-CNN, and Random Forest models on physics-derived, reagent-intensity, and temporal-lag features, producing an explainable signal for proactive reagent control.

AUTHORS	AFFILIATION	CLASSIFICATION
T. Davhana	**Axiom Process Technology**	**Public Release**
D. Mthimunye	**Nelspruit, South Africa**	**Vol 01, No 02**

ABSTRACT

Reverse cationic flotation is the dominant route for upgrading iron ore concentrates, with the silica content of the concentrate determining the saleable grade and the corresponding price realisation. The principal closed-loop measurement of SiO_2 grade is laboratory X-Ray Fluorescence assay on grab samples, with turnaround times of two to four hours from sampling to availability. The control implication is that all reagent adjustments are retrospective by definition. This paper presents an ensemble soft sensor architecture that predicts SiO_2 grade thirty minutes into the future from a continuously available SCADA feature vector. The architecture stacks an LSTM network for long-range reagent-residence memory, a one-dimensional Convolutional Neural Network for local temporal patterns, and a Random Forest regressor for feature-importance transparency. A ridge meta-learner produces a point estimate together with a ninety-five percent prediction interval. Approximately forty-five to sixty-five engineered features are organised into five categories: physics-derived flotation kinetics, reagent-intensity normalisations, feed characterisation, temporal lag windows reflecting reagent residence time, and statistical-window stability indicators. The sensor enables proactive amine, starch, wash-water, and froth-depth adjustments *before* the laboratory assay confirms the disturbance.

KEYWORDS

Soft sensor, iron ore, reverse cationic flotation, silica, LSTM, 1D CNN, Random Forest, ensemble stacking, prediction interval, proactive control.

1. Introduction

Iron ore reverse cationic flotation upgrades the concentrate by floating quartz and silicate gangue with an amine collector while depressing iron-bearing minerals with starch. The grade-recovery placement of the operating point is sensitive to the air flow per cell, froth depth, the dosage and ratio of amine to starch, the feed pH, and the hydrodynamic regime in the collection and froth zones. Each of these variables is continuously instrumented at the DCS level, yet the only direct measurement of the objective function, the silica content of the concentrate, arrives by laboratory XRF assay on the order of hours after the sample is drawn. The result is that the operator is steering a process whose principal output is observed retrospectively.

The cost of this latency is twofold. Grade departures that exceed contract specification attract penalties or rejection at the customer; grade overshoots above specification represent recovery sacrificed to a conservatism the operator has no faster way to unwind. The economic prize for a thirty-minute forward estimate of concentrate SiO_2 grade is the elimination of both ends of this distribution. A soft sensor that closes this gap must be continuously available, physically interpretable to the resident metallurgist, and explainable to the operator who acts on it.

2. Process and Data Architecture

2.1 SCADA input set

The architecture is fed by twelve continuously available SCADA tags at a one-second DCS scan rate. Per cell, the air flow rate is supplied by a rotameter or mass-flow transmitter; the froth depth or column level by an ultrasonic or float sensor. Feed flow rate is taken from a magnetic flow meter, feed density and percent solids from a nuclear density gauge, and feed pH from an inline electrode. Amine collector and starch depressant dosages are read from dosing-pump flow transmitters. Wash-water flow per column, concentrate flow, tailings flow, and slurry temperature complete the measurement set. Feed SiO_2 and iron grade from periodic XRF assays serve as training labels only and do not appear as real-time inputs.

2.2 Preprocessing

Raw SCADA streams are subjected to timestamp alignment, interquartile-range outlier removal at three sigma, forward-fill on laboratory channels to maintain a continuous training label, and min-max normalisation. The train, validation, and test split is strictly chronological. The test set is constructed to deliberately include ore transitions, ensuring that generalisation is evaluated against the regime changes the sensor must survive in production.

3. Feature Engineering

Five feature categories combine to a vector of approximately forty-five to sixty-five features, presented in parallel to each base learner.

3.1 Physics-derived features

Superficial gas velocity is computed as J_g = Q_air ÷ A_cell, governing bubble generation and gas holdup in the collection zone. Froth residence time is computed as τ_f = h_froth ÷ v_froth, governing drainage and entrainment behaviour in the froth phase. Hydraulic bias is computed as B = Q_wash –

Q_conc; a positive bias drives net downflow and is associated with cleaner iron concentrate.

3.2 Reagent-intensity features

Specific amine dosage D_a = Q_amine ÷ Q_feed and specific starch dosage D_s = Q_starch ÷ Q_feed normalise reagent consumption to feed rate in grams per ton. The selectivity index SI = D_a ÷ D_s captures the collector-depressant balance as a single dimensionless ratio and serves as a direct proxy for silica selectivity.

3.3 Feed characterisation features

Feed solids content is derived from the nuclear density gauge through the standard volumetric relation. Volumetric throughput Q_v = Q_feed ÷ ρ_slurry captures the actual volume flow into the cell and is the correct turbulence proxy in the collection zone. A pH-deviation feature ΔpH = pH – pH_target encodes the distance from the amine optimum, conventionally near pH 10.5, where selectivity is maximal.

3.4 Temporal lag features

Reagent residence time in reverse cationic flotation is approximately twenty to thirty minutes. The feature set therefore includes lagged values of air flow at five, ten, and twenty minutes; lagged amine dosage at ten, twenty, and thirty minutes; lagged froth depth at five, ten, and twenty minutes; and lagged feed density at ten, twenty, and thirty minutes. These lags provide the architecture with the memory it needs to attribute present grade outcomes to upstream reagent and ore changes.

3.5 Statistical window features

Fifteen-minute rolling means establish a trend baseline; fifteen-minute rolling standard deviations indicate process stability; an exponentially weighted moving average with λ = 0.1 supplies an exponentially decaying memory for slow drifts. Rate-of-change features ΔX ÷ Δt over a five-minute horizon function as leading indicators of ore transition.

4. Model Architecture

4.1 Base learners, parallel training

Three base learners are trained in parallel on the full feature vector. The LSTM network operates on sequences of thirty time steps and is intended to capture the long-range reagent-residence memory that the lag features expose. The 1D Convolutional Neural Network uses kernels of five to fifteen steps and is intended to identify local temporal patterns in the SCADA streams. The Random Forest regressor uses five hundred trees and contributes feature-importance ranking that makes the contribution of each engineered feature transparent to the metallurgist.

4.2 Ensemble stacking layer

The three base-learner outputs are combined through a ridge meta-learner that produces both a point estimate and a ninety-five percent prediction interval. The ridge weights are updated on a rolling validation window so that the relative confidence in each base learner adapts to the regime the plant is currently in. Where the LSTM and 1D CNN disagree sharply, the prediction interval widens; the metallurgist is alerted that the sensor itself is uncertain, not given a single number with no context.

4.3 Output and control integration

The output is the predicted concentrate SiO_2 grade at T+30 minutes, updated every DCS scan, accompanied by a ninety-five percent prediction interval and the SHAP-style attribution from the Random Forest leg of the ensemble. The control layer above the sensor responds proactively: amine dosage is adjusted up or down, starch dosage is rebalanced, wash water is trimmed, and the froth-depth setpoint is shifted before the laboratory assay confirms the disturbance. The lab assay returns as a delayed training label, closing the supervisory loop, not the regulatory one.

5. Architectural Position in the Axiom Control Stack

The SiO_2 soft sensor occupies the third layer of the Axiom control hierarchy, alongside the ore-hardness soft sensor and other physics-informed estimators. It sits above a healthy DCS and above the anomaly-detection layer that protects the regulatory foundation from sensor dropout, sticky valves, and dead-banded transmitters. The sensor output enters the Advanced Process Control layer as a feedforward signal, where the hybrid FLC and MPC stack converts predicted grade departures into reagent and air-flow moves. The AI Optimisation Layer above the APC consumes both the prediction and the prediction interval when computing the thirty-minute economic setpoint recommendations that the metallurgist accepts or rejects.

6. Conclusion

The laboratory turnaround on concentrate SiO_2 assay is the binding latency in iron ore reverse flotation control. The ensemble soft sensor presented in this paper closes that gap by combining physics-derived, reagent-intensity, feed-characterisation, temporal-lag, and statistical-window features through a stacked LSTM, 1D CNN, and Random Forest architecture. The output is a continuously available, prediction-interval-qualified, feature-attributable forecast of concentrate grade thirty minutes ahead. The control consequence is the elimination of both grade penalties and recovery sacrifices attributable to the assay latency.

References

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), pp. 1735–1780.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), pp. 5–32.

Lundberg, S.M. and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, pp. 4765–4774.

Wills, B.A. and Finch, J.A. (2016). *Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery*, 8th ed. Butterworth-Heinemann, Oxford.