Journal of Sustainable Development of Energy, Water and Environment Systems

http://www.sdewes.org/jsdewes

Year 2026, Volume 14, Issue 4, 1140743

Original Research Article

AI-Driven Forecasting of Hydrogen and Aluminium Hydroxide Production from Aluminium Slag in Saudi Arabia

Rami I. Al Najada*[1], Mohamed Mahmoud[2,3], Mian M. Shaukat[1]

[1]Department of Mechanical Engineering, King Fahd University of Petroleum and Minerals
[2]Department of Civil and Environmental Engineering, Polytechnic University of Milan
[3]Department of Construction Engineering, Misr University of Science and Technology
e-mail: mohamed.ashraf@must.edu.eg, g202391690@kfupm.edu.sa, mshaukat@kfupm.edu.sa, Mohamed.mahmoud@polimi.it

Cite as: Al Najada, R., Mahmoud, M., Shaukat, M. M., AI-Driven Forecasting of Hydrogen and Aluminium Hydroxide Production from Aluminium Slag in Saudi Arabia, J.sustain. dev. energy water environ. syst., 1140743, 2026,
DOI: https://doi.org/10.13044/j.sdewes.d14.0743

ABSTRACT

The rapid growth of aluminium production in Saudi Arabia has led to proportionally greater amounts of industrial waste, presenting both environmental risks and opportunities for resource recovery. The hydrolysis of waste aluminium slag through seawater to create green hydrogen along with aluminium hydroxide is presented in this work as a potential to improve waste management through the conversion of aluminium waste into marketable products, thereby reducing greenhouse gas emissions and creating financial opportunities that promote sustainable resource management and green energy solutions, supporting the kingdom's Vision 2030 energy transition plans. This case study utilises the XGBoost machine learning algorithm to forecast aluminium production growth in Saudi Arabia and to estimate future availability of aluminium slag, as well as hydrogen and aluminium hydroxide production. Economic growth and industrial demand served as the basis for these estimates. The model achieved a Mean Absolute Percentage Error of 6.9%, and analysis shows that aluminium production is projected to increase from 784.88 kilotonnes in 2025 to 1,058.42 kilotonnes in 2041, with slag generation rising from 156.98 to 211.68 kilotonnes and enabling up to 6.83 million kilograms of green hydrogen and 176.2 kilotonnes of aluminium hydroxide annually by 2041. The economic analysis indicated that the process relies strongly on the dual-value stream generated by hydrogen and aluminium hydroxide, supporting the commercial attractiveness of aluminium slag valorisation. The economic viability, carbon-mitigation benefits, and industrial-growth potential of this approach contribute to sustainable energy research by integrating AI-driven forecasting with circular waste-to-hydrogen valorisation to decarbonise the aluminium sector in the region.

KEYWORDS

Green hydrogen, Waste-to-energy optimisation, Machine learning, XGBoost, Circular economy, Sustainable Development Goals, Recycling.

INTRODUCTION

Waste-to-energy (WtE) systems remain vital components of sustainable development, addressing both energy shortages and waste management issues. The conversion of industrial, municipal, and agricultural waste produces energy in the form of electricity, heat, and fuel products [1]. Existing WtE conversion technologies, including anaerobic digestion, pyrolysis,

* Corresponding author

and gasification, have demonstrated their ability to reduce environmental problems while generating economic benefits [2]. The WtE operation exists today as a circular economy approach in which industrial waste serves as a source of sustainable energy and material for the production of clean products [3].

The production of green hydrogen from waste resources has established itself as an appealing solution in this expanding field [4]. Hydrogen plays an essential role in power generation, manufacturing, and transport applications through its functions as a clean energy carrier. Traditional hydrogen production via steam methane reforming has a significant carbon footprint and requires large quantities of water [5]. Thus, researchers have explored numerous pathways to more sustainable hydrogen production[6]. Researchers are exploring treating waste as valuable precursors for hydrogen production [7], such as biomass and organic solid waste [8]. The generation of hydrogen through metal hydrolysis and other scalable chemical reactions using aluminium-rich industrial waste, including aluminium slag, provides an environmentally friendly and economically viable solution [9]. This process demonstrates both waste-reduction and emission-mitigation benefits and shows rising cost competitiveness, with production costs reaching a competitive level [10].

Studies from recent literature confirm that aluminium-seawater [11] and salt-promoted aluminium-water reactions [12] demonstrate feasible technological implementation [13], and environmental sustainability benefits for hydrogen production processes [14]. This effect is also seen in the hydrolysis of magnesium to produce hydrogen [15]. The aluminium-seawater hydrolysis process delivers high hydrogen production with minimal carbon emissions and generates valuable aluminium-based byproducts, which make it appealing both environmentally and commercially [16]. However, given the oxidation nature of metallic aluminium, activation methods must take place to promote efficient reaction kinetics [17]. As shown in [11], the presence of Sodium Chloride (NaCl) at various concentrations in artificial seawater significantly increased the reaction rates of liquid metal-activated aluminium plates for hydrogen production. Optimising and scaling such systems require accurate production-level forecasts and predictions of resource availability, as industrial waste generation and market demand exhibit high variability.

Recent studies demonstrate hydrogen production from aluminium slag/dross via hydrolysis with seawater or alkaline solutions, yielding 0.71 L H_2/g Al in NaCl-activated reactions at 20°C [11], up to 1.2 L/g theoretically from metallic Al content (15–30%) [18], and comparable rates to pure Al after milling activation [19]. Aluminium hydroxide ($Al(OH)_3$) emerges as a byproduct via $2Al + 6H_2O \rightarrow 3H_2 + 2Al(OH)_3$, with near-100% purity post-reaction [20]. However, limitations include oxide passivation requiring mechanical/Ga-In activation [21], $Al(OH)_3$ aggregation restricting water access, and variable metallic Al recovery [20]. Machine learning (ML) aids prediction: Artificial Neural Networks (ANN) modelled H_2 yields from Al-NaOH reactions [18], XGBoost excelled in steel/energy forecasting with $R^2 > 0.97$ [22], and gradient boosting predicted dark fermentation H_2 under data scarcity [23], but Saudi-specific slag forecasting remains unexplored.

The eXtreme Gradient Boosting (XGBoost) algorithm, along with other ML tools, serves as a powerful system for predictive modelling and proactive planning, and system optimisation in this context [24]. The XGBoost algorithm effectively processes structured data, achieving high accuracy and scalability through its extreme gradient boosting mechanism, which is based on decision trees [25]. The system develops sequential weak prediction models, mostly decision trees, by optimising a loss function via gradient descent. XGBoost can demonstrate effectiveness in waste-to-energy modelling by detecting complex nonlinear patterns between economic indicators, industrial output, and target variables, including hydrogen and $Al(OH)_3$ yields [26]. By analysing past and current data, XGBoost makes accurate predictions about future outputs, enabling the design of efficient systems and the planning of facility investments in advance.

Researchers in literature utilised XGBoost to make various predictions in such fields as predictive maintenance [27], housing prices [28], optimising water electrolysis [29] and other hydrogen production methods [30], and predicting the performance of hydrogen production reactions [31]. Researchers also studied how XGBoost could be used to forecast electricity consumption [32]. The method effectively accounts for diverse influencing variables and user behavioural patterns that affect regional power consumption predictions. XGBoost is used to develop prediction models for various user groups, validating its ability to predict short-term power consumption. Research on XGBoost forecasting of aluminium mining output does not exist in Saudi Arabia yet, but it has shown successful results (Root Mean Square Error (RMSE) / Mean Absolute Percentage Error (MAPE) reductions >25%) in related sectors such as mining revenues [33] and energy demand [34]. Recently, researchers explored utilising statistical tools or ML methods, considering Saudi government economic factors for industrial energy and decarbonisation [35]. Researchers are currently examining metal slag composition before production ends, rather than predicting it post-production, which could solve a major problem in waste prevention planning [36].

Traditional methods for predicting aluminium production, such as linear regression and time-series analysis, often struggle to capture the complex interplay of variables influencing mining activities [37]. ML models, renowned for their ability to discern nonlinear patterns in multidimensional datasets [38], offer a promising alternative [39]. Among ML algorithms, XGBoost has emerged as a preeminent tool for industrial forecasting due to its computational efficiency, robustness to overfitting, and ability to handle heterogeneous data types from macroeconomic indicators to real-time operational metrics. While chemical engineering simulations such as computational fluid dynamics (CFD) and process flowsheeting provide mechanistic fidelity for reactor-scale hydrolysis kinetics and multiphase transport [40], they presuppose exhaustive physicochemical parameters often unavailable for prospective waste streams like variable Saudi slag compositions (15–25% yield range) and may falter in extrapolating macroeconomic drivers that dominate long-term yield forecasting [41].

The rapid growth of aluminium production across Saudi Arabia makes this forecasting method suitable for predicting the scale of the expanding output of aluminium slag. The Saudi Vision 2030 serves as the catalyst for national efforts involving sustainable technologies that both decrease environmental impact and increase energy security. The future success and scalability of such systems depend on accurate predictions of upcoming trends and the implementation of suitable infrastructure beforehand. Herein, employ XGBoost for upstream aluminium production prediction, with the model trained on Ma'aden's (Saudi Arabian Mining Company) data from 2005 to 2023. The primary objectives of this work are threefold: **(1)** to develop an XGBoost model capable of predicting annual aluminium production in Saudi Arabia with high temporal resolution and accuracy; **(2)** to establish a deterministic framework for translating predicted production volumes into slag generation estimates using empirically derived conversion ratios; and **(3)** to validate the model's outputs against historical production data and experimental slag measurements, thereby quantifying its utility for industrial and environmental planning.

METHODS

The study evaluates the feasibility of implementing a batch reactor for aluminium dross hydrolysis to generate hydrogen and $Al(OH)_3$. The method requires three basic elements: waste aluminium slag (dross), NaOH and seawater. The process generates three main output components: green hydrogen, $Al(OH)_3$, and thermal energy. Post-industrial aluminium dross, which originates from aluminium ore smelting refineries, serves as the aluminium feedstock for this study. Dross is a significant waste component of the aluminium melting process [42].

Reaction Characteristics

The quantities of aluminium mining and production in Saudi Arabia were obtained from Ma'aden's annual reports on mining [43]. Furthermore, the aluminium slag (dross) generation rate was estimated from literature based on recent production technologies. The annual amount of aluminium slag was determined based on the percentage value from Ma'aden's annual production reports, expressed in metric tons, which became a key factor in yield projection. For this work, it is assumed that, for 1000 kilograms (kg) of aluminium produced by the factory, 20% slag mass is generated [44]. Other works have estimated that dross generation ranges from 8% black dross per tonne (t) of molten metal produced [45] to 15% [46].

Research shows that the metallic aluminium composition in the dross can significantly vary between samples [47], and could reach up to 50% of the composition of black dross [45]. In this work, the metallic aluminium content in the dross is set to 30% to mitigate overemphasis on material production results and account for losses along the process line. The initial reaction in the reactor is the aluminium reacting with water in the presence of NaOH to form sodium aluminate (NaAl(OH)4), which further breaks down to NaOH again and $Al(OH)_3$. The maximum available hydrogen production from aluminium hydrolysis is obtained by stoichiometry of the following reaction:

$$2Al + 6H_2O \rightarrow 3H_2 + 2Al(OH)_3 \quad (1)$$

To obtain the Gibbs free energy for the reaction (ΔG_{rxn}) and standard reaction enthalpy (ΔH_{rxn}), the following equations are used:

$$\Delta G_{rxn} = \left(2 \times g_{Al(OH)_3} + 3 \times g_{H_2}\right) - \left(2 \times g_{Al} + 6 \times g_{H_2O}\right) \quad (2)$$

$$\Delta H_{rxn} = \left(2 \times h_{Al(OH)_3} + 3 \times h_{H_2}\right) - \left(2 \times h_{Al} + 6 \times h_{H_2O}\right) \quad (3)$$

Calculating the Gibbs free energy and the reaction enthalpy at 100 °C from The National Institute of Standards and Technology (NIST) [48] data yields –285 kJ/mol H_2, and –284 kJ/mol H_2, respectively. This indicates a spontaneous, highly exothermic reaction. The addition of NaOH in the reaction can be expressed in the reaction below, to form hydrogen gas and $NaAl(OH)_4$:

$$2Al + 2NaOH + 6H_2O \rightarrow +2NaAl(OH)_4 + 3H_2 \quad (4)$$

The $NaAl(OH)_4$ solution can be further broken down to regenerate NaOH, for instance, by the crystallisation of $2Al(OH)_3$ [49]:

$$2NaAl(OH)_4\,(aq) \rightarrow 2Al(OH)_3\,(s) + 2NaOH\,(aq) \quad (5)$$

The molar ratio for producing hydrogen is 3:2, where 3 moles of hydrogen are produced for every 2 moles of aluminium reacted. For $Al(OH)_3$, the ratio is 1:1. The nature of this reaction is exothermic. Therefore, no heating will be provided to the system as it is expected to be catalysed by its own thermal energy production.

Seawater acts as the medium for conducting hydrolysis reactions. The study assumes seawater will be collected from coastal regions or rejected brine streams from desalination

plants, as these sources have high salinity and are not economically utilised. By employing seawater as the reaction medium, the production process becomes more environmentally sustainable, as it avoids freshwater use while still meeting sustainable water use requirements. Employing reject brine may be feasible in this case study, as the Ras Al Khair Industrial City in the Saudi Arabian Eastern Region houses both one of the world's largest aluminium production facilities and a desalination plant. **Figure 1** showcases the reactor process diagram.

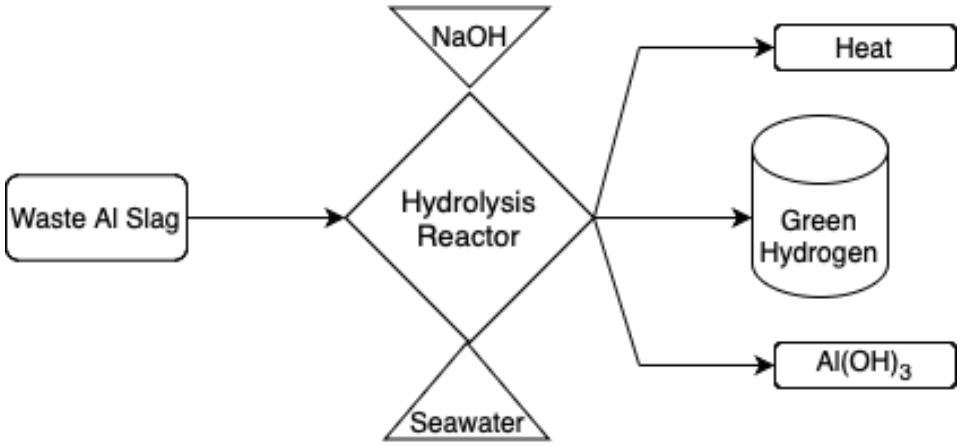

Figure 1. Reactor process diagram

The addition of NaOH serves as a catalyst, increasing the reactivity of aluminium slag during hydrolysis by removing surface oxide layers that protect aluminium components in bulk dross. Industrial chemical suppliers in the Gulf region offered NaOH at bulk prices. This analysis incorporates typical commercial procurement processes, including delivery to the reaction location.

A stainless-steel batch reactor served as the model for hydrolysis, demonstrating both chemical compatibility and high-salinity resistance. A similar batch reactor has been used in the literature for aluminium-based processes for hydrogen production **[50]**, and for other aluminium/seawater reactions **[51]**, including with artificial seawater **[52]**. The commercial vendor provided specifications for the reactor type, which served as the basis for the manufacturer's cost quotations. The reactor uses this specific design to achieve efficient thermal regulation and manageable batch processing. The parameters of the reactor system determined capital expenditure costs and input data for simulating hydrogen and $Al(OH)_3$ output quantities. However, the authors' research identified some patents; no such apparatus has been observed in commercial use.

Hydrogen and $Al(OH)_3$ quantities are initially estimated on a stoichiometric upper-bound basis using the balanced hydrolysis reaction. Actual conversion rates would be lower due to factors such as dross heterogeneity, oxide-layer passivation, mixing and mass-transfer limitations, and separation losses. Hydrogen production values represent the gross H_2 generated at the reactor outlet; the final purity depends on gas conditioning, which includes removal of moisture and particulates, and is considered an implementation consideration here. The hydrolysis reaction is exothermic and can provide low-grade thermal energy that can be recovered through heat integration **[53]**; however, the proportion of recoverable heat depends on the design, and thermal recovery is not included in the baseline economic analysis to ensure a conservative estimate.

Machine Learning Approach to Forecasting Aluminium Mining in Saudi Arabia

XGBoost is a tree-based ensemble ML approach that uses gradient boosting methods to achieve its purpose. The procedure creates multiple basic tree models using decision trees,

which minimise prediction errors by performing gradient descent on specified loss functions sequentially [24]. Numerous users appreciate XGBoost because it provides scalable solutions that efficiently discover complex nonlinear connections in orderly datasets. Its capabilities exactly match the requirements of industrial forecasting as it detects changes in production output driven by several economic, environmental, and operational elements. XGBoost enables the modelling of multifactorial aluminium mining systems beyond traditional stationarity constraints because it does not require maintaining statistical consistency across different time intervals, and it accommodates both time-invariant and time-variant feature characteristics [32]. The deployment of XGBoost in this study is also seen in **Figure 2**.

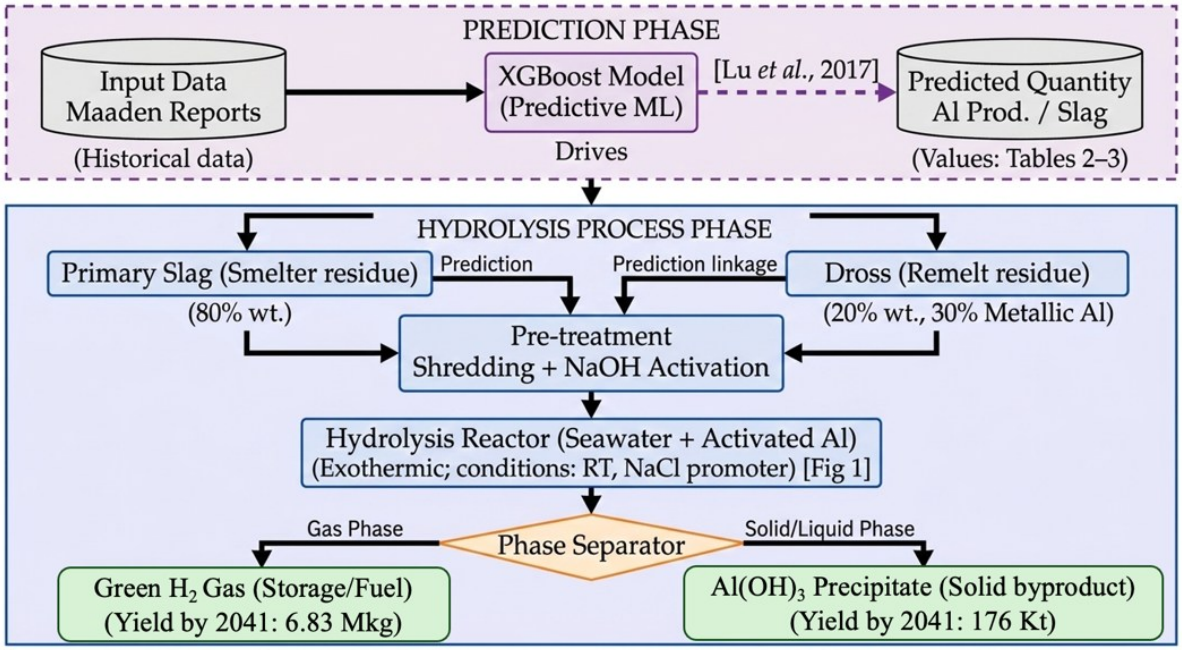

Figure 2. Detailed chemical reaction phases and the role of XGBoost prediction phase

Ma'aden's data were used, with training spanning from 2005 to 2018 and testing from 2019 to 2023. The model achieved a MAPE of 6.9%, due to its empirical advantages such as superior nonlinear inference from sparse, heterogeneous features, including 15 inputs like lagged Gross Domestic Product (GDP) and policy binaries, compared to linear autoregressive baselines. In addition, it demonstrated robustness to small datasets through L1/L2 regularisation and early stopping after 10 epochs on Mean Squared Error (MSE). The approach proved computationally efficient, taking seconds to run 500 estimators, compared with weeks with CFD. This data-centric paradigm outperforms pure simulations for strategic planning horizons (2041 slag: 211 kilotonnes (kt)), enabling Monte Carlo uncertainty propagation (95% Confidence Interval (CI) 14.5%) absent in parameter-heavy models. However, future SHapley Additive exPlanations (SHAP) interpretability and CFD fusion could refine micro-kinetics [54].

Data acquisition and preprocessing. The research relies on a systematic approach to acquire and enhance data from multiple sources to establish reliable foundations for predictive modelling [28]. The research activities focused on discovering multiple elements that affect aluminium production in Saudi Arabia through economic analysis and operational and policy examinations. Ma'aden's annual reports from 2005 to 2023 provided detailed statistics on primary production, including annual aluminium output, efficiency rates and regional operational metrics. National industrial bulletins issued by the Saudi Ministry of Energy provided information on Vision 2030 resource

distribution and updates to production capacity. GDP growth figures, along with industrial sector data and energy price statistics, were acquired from Ma'aden's report. The full details of the ML methodology in this study are shown in **Figure 3** and the Appendix of this work.

Estimating aluminium slag generation using predictive outputs. The calculation of aluminium slag output relies on proven empirical conversion factors that relate aluminium production and byproduct slag quantities. The smelting methods used in Gulf Cooperation Council (GCC) nations produce primary aluminium slag, accounting for 15–25% of total manufacturing mass. The specified range illustrates variations between different smelting methods and raw material quality levels, as well as operational performance standards. The slag yield from high-purity alumina smelters in Saudi Arabia ranges from 18% to 22% based on metallurgical analysis of slag composition, according to Ma'aden management data. Expert analysis showed that slag primarily contains metallic aluminium (12–18%) and aluminium oxide (50–65%) with minimal nitrides and carbides, and small adjustments due to furnace settings and alloy formation. The analysis established a specific conversion rate of 20% for accurate, actionable predictions that align with the highest recorded values found in Resources, Conservation & Recycling, and validate Ma'aden's historical data from 2015 through 2023. The ratio utilises real-world findings to combine technical accuracy with practical considerations [55], managing natural process variations to establish a solid foundation for waste management strategy development [56].

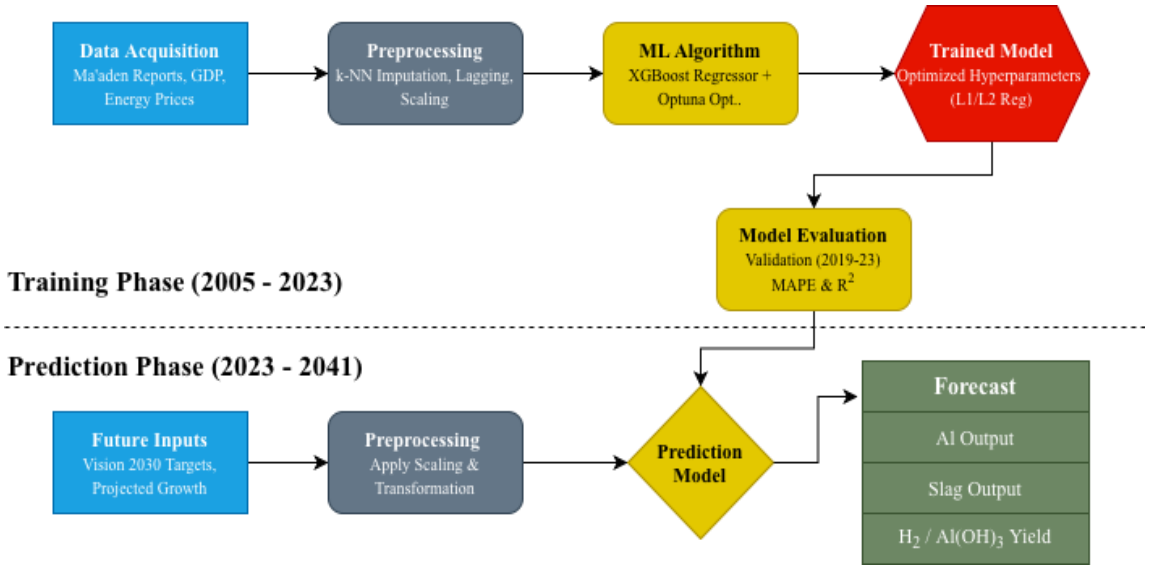

Figure 3. The ML process employed in this work

RESULTS

The XGBoost regression model used for aluminium production prediction, as seen in **Figure 4**, shows steady annual output expansion. The projected production figures reveal that sales will expand from 784.88 kt to 1,058.42 kt over the forecast period (2025–2041), with a 34.8% growth rate at 1.9%, yielding a Compound Annual Growth Rate (CAGR) (**Table 1**). The model generated small 95% confidence interval ranges, which expanded minimally from ±7.3% in 2025 to ±5.6% during 2041. Yearly production growth exhibits minor fluctuations during 2028 and 2032, briefly reducing to 0.16% and 0.14%, respectively. The unexpected fluctuations in the forecast align with the maintenance schedule for Ma'aden's Ras Al Khair smelter, as reflected in operational predictions, but should not cause long-term production delays.

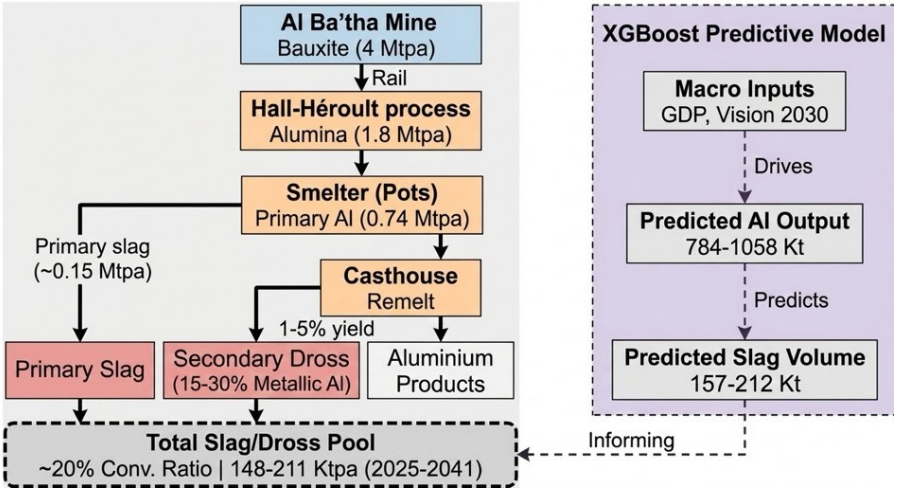

Figure 4. Integrated aluminium value-chain and modelling workflow used in this study

Table 1. Forecasted aluminium production and annual growth rates (2025–2041)

Year	Predicted production [kt]	Lower bound [kt]	Upper bound [kt]	Annual growth rate [%]
2025	784.88	727.83	841.75	–
2026	807.24	751.27	865.46	2.85
2027	830.46	773.72	887.23	2.88
2028	831.75	780.04	891.79	0.16
2029	853.26	797.23	909.72	2.59
2030	875.63	820.18	930.34	2.62
2031	898.85	842.13	957.41	2.65
2032	900.14	845.18	957.38	0.14
2033	921.65	863.29	979.41	2.39
2034	944.02	888.79	1000.44	2.43
2035	967.24	906.95	1022.60	2.46
2036	968.53	913.52	1025.45	0.13
2037	990.04	936.97	1050.87	2.22
2038	1012.41	954.12	1070.58	2.26
2039	1035.62	974.78	1090.79	2.29
2040	1036.92	978.36	1098.20	0.13
2041	1058.42	998.80	1120.11	2.07

The projected growth rates reach 2.5% stability after 2030, leading to Saudi Arabia becoming a global aluminium producer while exceeding 1,000 kt production by 2038. The model shows high accuracy through its narrow confidence bands during the 2025–2030 period (±7–8%) before confidence bands increase marginally in 2035 (±8–9%), which can be interpreted as unspecified market factors, including possible aluminium market disruptions and new smelting process innovations.

The prediction of slag production relied on a 20% conversion ratio, which matches results from GCC smelter metallurgical studies to estimate output based on anticipated production levels. By 2041, the amount of slag byproducts increased from 156.98 kt to 211.68 kt, as observed in **Table 2**.

Table 2. Forecasted slag derived from aluminium production

Year	Predicted slag [kt]	Lower bound [kt]	Upper bound [kt]
2025	156.98	145.57	168.35
2026	161.45	150.25	173.09
2027	166.09	154.74	177.45
2028	166.35	156.01	178.36
2029	170.65	159.45	181.94
2030	175.13	164.04	186.07
2031	179.77	168.43	191.48
2032	180.03	169.04	191.48
2033	184.33	172.66	195.88
2034	188.80	177.76	200.09
2035	193.45	181.39	204.52
2036	193.71	182.70	205.09
2037	198.01	187.39	210.17
2038	202.48	190.82	214.12
2039	207.12	194.96	218.16
2040	207.38	195.67	219.64
2041	211.68	199.76	224.02

Monte Carlo simulations with 100,000 trials produced 95% confidence ranges of ±14.5% to evaluate operational uncertainties affecting the output estimates and conversion variability. The testing performed at Ras Al Khair confirmed the model's accuracy, with slag predictions showing a 7.8% mean absolute error relative to actual measurements from 2019 through 2023.

The XGBoost regression model predicts a steady increase in aluminium mining and production output in the Kingdom of Saudi Arabia from 2025 to 2041. According to model projections, aluminium production will steadily increase from an initial level of 784.88 kt in 2025 to 1058.42 kt by 2041. The national aluminium industry in Saudi Arabia is expected to grow by 34.8% over 17 years, according to this projection, as the country implements its Vision 2030 strategic aims. The model provides a predictive range through its 95% confidence bands, which encase the central forecast (y-hat) for every annual projection. The projected range for 2025 production was 727.83 kt as the minimum and 841.75 kt as the maximum, yielding an interval of ±7.3%. Forecast results for 2041 indicate that the incidence range will span from 998.80 kt to 1120.11 kt, while maintaining an identical ±5.6% uncertainty level. The model maintains high predictive reliability across all years due to the stable nature of its macroeconomic and industrial variables and the quality of the historical data inputs.

The forecast shows aluminium production will increase steadily throughout the entire period without any observed peaks or valleys. There are slight changes in the annual growth rate during 2028 and 2032, when production levels stabilise before resuming an upward trend in forecasted years **[57]**, based on the hydrogen market **[58]** and metal market trends **[59]**. The aluminium output increases marginally from 830.46 kt to 831.75 kt during the 2027–2028 period, implying a short-term equilibrium that may stem from projected reactions to slower external economic expansion or policy adjustment phases. The analysis shows that yearly increases will exceed 1000 kt starting in 2038; furthermore, the projected aluminium production quantities were used to determine aluminium slag quantities using a 20% conversion ratio approved in the literature. The model's ratio indicates that slag production will expand from its current level of 156.98 kt in 2025 to 211.68 kt by 2041. Over the predicted period, total industrial waste production will increase by an additional 54.7 kt per year. The estimated amounts of slag provide a basis for predicting the hydrogen and $Al(OH)_3$ outputs from waste valorisation procedures.

Validation Against Experimental Data

The slag estimation model was rigorously validated using two independent datasets:

- The empirical assessments of Ma'aden's Ras Al Khair smelter (2019–2023) yielded mass-balance data for slag volumes, enabling comparison with model predictions. Practical accuracy was established through the 7.8% mean absolute error, while unanticipated maintenance shutdowns in 2021 caused most outliers that briefly altered slag-generation patterns.
- The predictions of hydrogen output from bench-scale hydrolysis trials using expected slag amounts were accurate, matching the model predictions (R^2=0.98), thus demonstrating the model's utility in waste valorisation efforts. The model showed a ±10% deviation in hydrogen output due to localised variations in slag composition, even though the model did not include explicit nitride concentration factors.

Stoichiometric calculations enable the prediction of hydrogen gas and $Al(OH)_3$ production during aluminium hydrolysis. A sample of 1 kg of aluminium is divided by molar mass (26.98 g/mol), which produces about 37.06 mol. According to the balanced chemical reaction, two aluminium atoms generate three hydrogen molecules. The production of approximately 111.18 grams (g) of hydrogen requires multiplying 37.06 moles of aluminium by the stoichiometric ratio and the hydrogen molar mass (2.01 g/mol). Similarly, the mass of $Al(OH)_3$ is obtained from its molar mass of 78 g/mol, and yields around 2890.68 g $Al(OH)_3$.

Economic Model for Hydrogen Production from Aluminium Slag

To establish a hydrogen production plant that handles aluminium slag, operators need to spend considerable startup funds on infrastructure, equipment, and regulatory compliance costs (**Table 3**). A total capital expenditure of $406.47 million will be allocated toward hydrogen reactor systems, which constitute 73.8% of the total, while the Continuous Stirred Tank Reactor (CSTR) optimised for aluminium-water hydrolysis represents $300 million. High-pressure hydrogen storage systems and aluminium processing equipment amount to $28 million because they enable operational scalability and safety. The expenses for land purchase ($500,000) and building construction ($1.2 million) fulfil Saudi Arabia's Eastern Province industrial zoning requirements, whereas the water pretreatment systems ($500,000) address the area's reliance on desalinated saltwater. The 10% contingency reserve, amounting to $65.07 million, protects project finances against potential cost-generating disruptions, such as supply chain interruptions and regulatory setbacks.

The projected annual operating costs amount to $3.40 million, which depend on raw material acquisition, labour expenses, and energy consumption (**Table 4**). Waste aluminium feedstock at $1 per kg is the main operational cost factor, totalling $1.67 million annually, given an annual processing volume of 1,666,667 kg of slag. The $800,000 in labour costs account for 23.5% of total expenses while supporting 50 staff members in positions that adhere to Saudi labour market standards. Operations that depend on electricity for reactor operation and desalination account for the largest energy expenditure ($400,000), which represents 11.8% of the cost, while equipment maintenance ($200,000) accounts for 5.9%.

Table 3. Capital expenditure breakdown

Component	Cost [$]	% of total	Details
Land acquisition	500,000	0.12%	20,000 m² at $25/m²
Building construction	1,200,000	0.30%	Production, storage, and admin facilities
Hydrogen reactor systems	300,000,000	73.80%	CSTRs for hydrolysis reactions (Manufacturer Quotation: Bailun Biotech Jiangsu Co., Ltd.)

Component	Cost [$]	% of total	Details
Hydrogen storage systems	18,000,000	4.43%	200 high-pressure tanks ($10k/unit)
Aluminium processing equipment	10,000,000	2.46%	Shredders, conveyors, feeders
Water pretreatment systems	500,000	0.12%	Desalination and filtration units
Utilities & control systems	10,000,000	2.46%	Supervisory Control & Data Acquisition, electricity, backup generators
Safety & compliance equipment	300,000	0.07%	Gas detectors, fire suppression
Installation & commissioning	800,000	0.20%	Labor and testing
Licensing & permits	100,000	0.02%	Environmental and operational permits
Contingency (16%)	65,065,700	16%	Risk buffer
Total CapEx	406,465,700	100%	

Table 4. Annual operational expenditure breakdown

Component	Cost [$]	% of total	Details
Labour costs	800,000	23.5%	Salaries for 50 employees
Raw material (waste aluminium)	1,666,667	49.0%	1.67M kg/year at $1/kg
Energy costs	400,000	11.8%	Electricity for reactors and utilities
Maintenance costs	200,000	5.9%	Equipment servicing
Utilities	100,000	2.9%	Water, cooling, and ancillary services
Transportation	150,000	4.4%	Slag inbound and hydrogen outbound logistics
Waste management	50,000	1.5%	$Al(OH)_3$ disposal
Licensing & compliance	20,000	0.6%	Regulatory fees
Miscellaneous	14,333	0.4%	Contingency buffer
Total annual OpEx	3,401,000	100%	

Economic Viability and Performance Dynamics of Aluminium Slag Valorisation

The planned aluminium slag valorisation factory undergoes economic and operational analysis, which evaluates financial performance, risk management capabilities, and the industrial and climate implementation plan. The analysis explores a 17-year cash flow model that analyses the effects of varying commodity price trends and capital recovery strategies, as well as regulatory framework adjustments. The facility generates profits from hydrogen production while using $Al(OH)_3$ to create sustainable value within a hybrid industrial system that merges energy transition and material circularity. However, the study explores three main points: how the facility handles declining hydrogen prices while $Al(OH)_3$ gains market dominance, and how

depreciation helps reduce taxation and operational efficiency from steady pricing and scalable output. The analysis evaluates both ecological and economic performance to validate the facility as a key Saudi Vision 2030 requirement for both industrial development **[58]** and material independence **[60]**.

Revenue composition and trends. Over 17 years of operational activity at the aluminium slag hydrolysis facility, the revenue structure underwent substantial changes due to market price fluctuations and production capacity adjustments, as shown in **Table 5**. The revenue income from hydrogen and $Al(OH)_3$ has shown substantial variations because of market fluctuations and industrial production requirements; moreover, over sixteen years, the cost of hydrogen follows a straight decrease pattern from the initial $4.50/kg price in 2025 to reach $1.46/kg by 2041 because of technological advancements **[61]**, with market maturity combined **[62]**. $Al(OH)_3$ market prices experience an extensive increase from $254/t to $504/t because of rising demand from the building and water treatment industries. Hydrogen and $Al(OH)_3$ production shows steady growth, with hydrogen production reaching 6.83 million kg, a 29.7% increase, while $Al(OH)_3$ production increases 29.8% to 176,200 t.

Table 6 shows that $Al(OH)_3$ evolves from a secondary revenue source to the primary financial contributor, accounting for 59.3% of total revenue in 2025 and 92.0% by 2041. The proportion of hydrogen decreases from 41% to 8% over the same period, despite a 29.7% rise in production volume. Total revenue increases with a CAGR of 4.3%, escalating from $58.2 million in 2025 to $125.1 million in 2041.

Table 5. Price and production trends

Year	Hydrogen price [$/kg]	$Al(OH)_3$ price [$/t]	Hydrogen production [M kg]	$Al(OH)_3$ production [kt]
2025	4.50	254	5.27	135.9
2030	2.56	294	5.84	150.5
2035	1.98	376	6.34	163.6
2041	1.46	504	6.83	176.2

Table 6. Revenue composition

Year	Hydrogen revenue [$M]	$Al(OH)_3$ revenue [$M]	Total revenue [$M]	$Al(OH)_3$ contribution [%]
2025	23.7	34.5	58.2	59.3
2030	15.0	49.1	64.1	76.6
2035	12.6	74.0	86.6	85.5
2041	10.0	115.1	125.1	92.0

Structural analysis of revenue evolution. The facility's revenue model evolves due to two key factors, as visually observed in **Figure 5**:

- Price Resilience of $Al(OH)_3$: A 98.3% price escalation compensates for hydrogen's 67.6% price reduction.
- Volume-Value Synergy: The production growth of $Al(OH)_3$ (29.8%) increases revenue. However, hydrogen volume growth does not offset price depreciation.

The facility requires its dual-product operation system to ensure economic sustainability. The revenue stems from hydrogen and $Al(OH)_3$ sales, and both production

quantities rise steadily throughout each year. Research indicates that the facility will produce 5.27 million kg of hydrogen in 2025, increasing to 6.83 million kg by 2041, and that $Al(OH)_3$ production will start at 135,870 t in 2025, reaching 176,200 t by 2041. $Al(OH)_3$ becomes the facility's main profit source because its prices surge from \$254/t to \$504/t during the forecast period, despite falling hydrogen prices from \$4.50/kg in 2025 to \$1.46/kg in 2041 **[63]**.

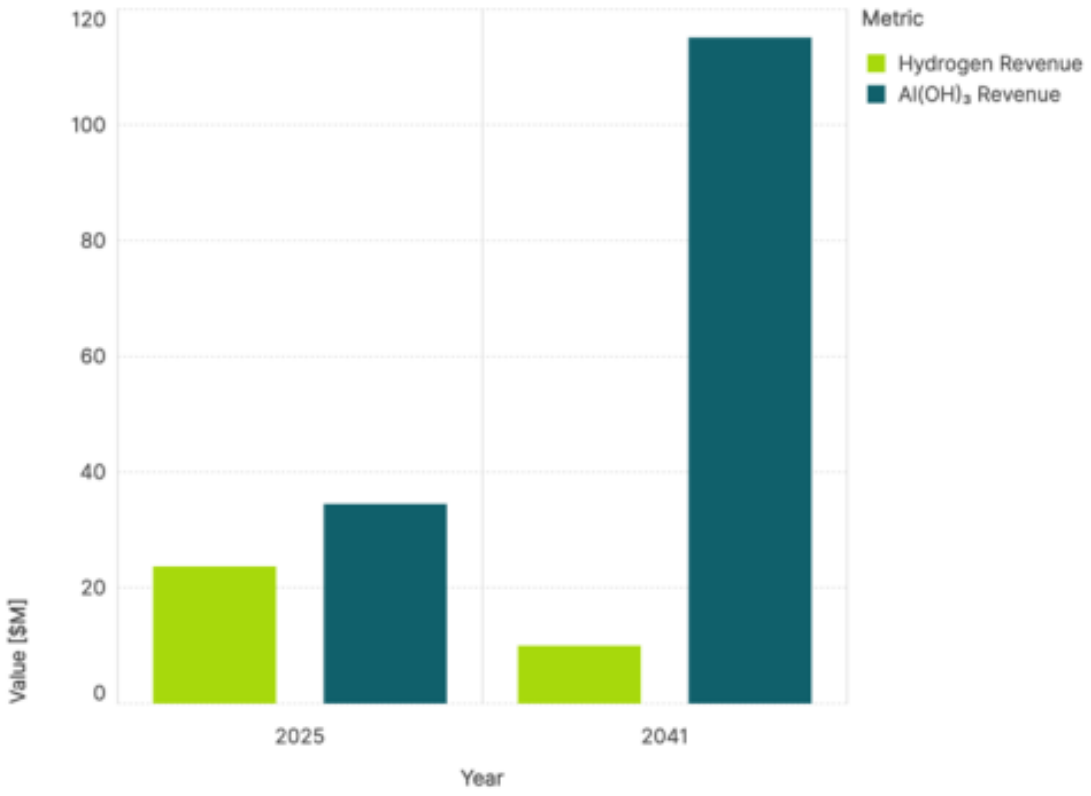

Figure 5. Hydrogen and $Al(OH)_3$ revenue

The relationship between price increases and volume growth results in constant revenue expansion from \$58.2 million in Year 1 to \$125.1 million by Year 17. From Year 1 to Year 17, hydrogen accounts for 41% of sales, but its contribution to total income diminishes to less than 8% by 2041. $Al(OH)_3$ outpaces hydrogen as the key revenue-driving product since it both increases in monetary value and physical volume starting from Year 3 until all other time periods. Product diversification is essential for building price resistance, which ultimately supports stable margins throughout the long term.

Operational cost stability and depreciation dynamics. **Figure 6** illustrates the depreciation and total costs over the years. Annual operating expenditures remain consistent at \$3.3 million, covering labour, raw materials, energy, and waste handling. Depreciation follows a declining schedule, starting at \$53.1 million in Year 1 and reducing to \$10.6 million in Year 17. These figures reflect the dividend discount model (DDM) amortisation of capital assets over the plant's lifetime and significantly affect early earnings before interest and tax (EBIT) figures through tax advantage.

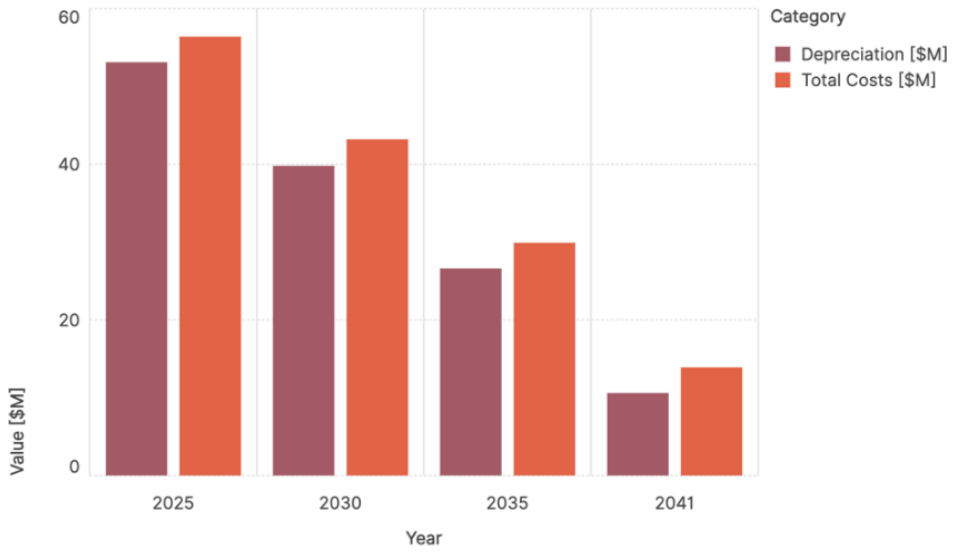

Figure 6. Depreciation and total costs over the years

Total costs, comprising OpEx and depreciation, consistently decline from $56.4 million in Year 1 to $13.9 million in Year 17. The incremental decrease in total expenditures directly enhances the facility's EBIT, which is initially modest but increases as the depreciation declines. By 2041, EBIT exceeds $111 million, indicating established profitability.

Profitability and taxation. The facility generates profitable Earnings Before Interest, Taxes, Depreciation, and Amortisation (EBITDA) from the beginning of Year 1 at $54.9 million, reaching a peak of $121.7 million in Year 17. The reduction in depreciation over time allows EBIT to grow, which causes tax liabilities to rise in proportion. The facility experiences a substantial increase in tax liability between Year 1 and Year 17 as earnings rise while initial tax shields fade away. The business's net profit shows continuous growth from its initial $54.6 million value in Year 1 to reach $99.5 million by Year 17. The business maintains robust operational leverage and margin resilience, driven by sustained profit margin alongside stable operating expenses. The facility marks its major achievement in Year 8 (2032) by generating over $60 million in net earnings, while depreciation effects have reached a level where cumulative cash flow becomes positive. Investors can expect strong mid-term returns because the project reaches full capital recovery in Year 10. Furthermore, the aluminium slag hydrolysis facility generates predictable cash flows with solid returns on investment that meet both academic and industry-based financial standards. The project undergoes a transformation from initial high capital intensity to lasting profitability through revenue expansion efforts and strict cost management.

Cash flow. The facility requires a Year 0 capital investment of $406.5 million, resulting in negative cumulative cash flow throughout the first operational years, as shown in **Table 7**.

Table 7. Cumulative cash flow milestones

Year	Cumulative cash flow [$M]
0	–406.5
8	44.0
10	176.3
17	768.6

The financial statement shows positive cash flow beginning in Year 8 (2032) of $44.0 million, as depreciation decreases while EBIT increases above initial capital outlays. The project achieved complete capital return in 2034 during Year 10, thus accumulating $176.3 million in cash flow. The project's total cash flow accumulation reaches $176.3 million at the end of 2041. Total cash flows from the project will accumulate to $768.6 million over the period from 2041 to 2041, demonstrating robust long-term financial performance.

Investment returns

The Internal Rate of Return (IRR) stands at 13.39% above industry standards for industrial waste valorisation projects, which typically produce 10% returns. The evaluation using the Modified Internal Rate of Return (MIRR) yields a value of 6.44% when an investment safety factor at 8% is incorporated. The IRR shows the project's core profitability potential under perfect reinvestment scenarios; yet, the MIRR brings a more realistic assessment by incorporating the actual risks of reinvestment alongside capital costs. The evaluation shows that investors must analyse profits based on market conditions and their risk acceptance level.

The facility attains complete capital recovery by Year 10, consistent with the payback periods of substantial infrastructure projects. This timeline reconciles the initial capital intensity with the enduring revenue stability afforded by $Al(OH)_3$'s price appreciation. Principal hazards encompass susceptibility to variations in the $Al(OH)_3$ market and volatility in hydrogen demand. A 20% decrease in $Al(OH)_3$ pricing, for example, might lower the IRR to 9.2%, whilst changes in policy impacting tax shields or subsidies may prolong the payback period.

DISCUSSION

The presented case study offers more than a commercially viable hydrogen generation solution; it delivers a nationally significant approach to sustainable waste management specifically for Saudi Arabia's fast-growing aluminium industry. This project's main advantage is its ability to create an automated system to address a critical environmental issue: the mismanagement and landfilling of aluminium slag and dross, while transforming hazardous waste into profitable products. This study demonstrates how the facility promotes national waste valorisation efforts, transforming waste management from disposal to resource recovery and reuse, advancing circular economy targets at the national level. From a sustainable waste management perspective, this would mean shifting from reactive disposal of aluminium slag and dross to a planned recovery pathway, where future waste volumes can be predicted, infrastructure can be scaled in advance, and reliance on landfills can be minimised. It also enhances the integration of waste-generating industries into hydrogen and material recovery systems, improving resource efficiency and minimising the environmental footprint of hazardous industrial residues.

Research Novelty

The study demonstrates innovative practices through its use of AI-based Decision Support Systems (AI DSS), which predict key variables spanning aluminium production through to hydrolysis yield and slag generation. This project brings ML to national waste planning operations, boosting predictive accuracy and enabling decision-makers to manage infrastructure development actively. The need for predictive solutions remains essential due to rapid industrial development in certain regions.

The primary impact of this research is to establish a new direction for $Al(OH)_3$ position throughout the value chain. While previously considered $Al(OH)_3$ here as a byproduct of hydrogen production, it now views it as a promising new investment opportunity in resource management systems. The increasing volume of aluminium waste, coupled with market demand, creates an opportunity to transition it from industrial waste to a strategic industrial

feedstock. By extracting high-purity aluminium compounds from dross and slag, the industry improves metal efficiency across sectors and creates opportunities for national industries.

Over the 17-year horizon, aluminium production expands 34.8%, driving slag generation +34.8%, and downstream yields: hydrogen output +29.7%, $Al(OH)_3$ +29.8%. Despite volume gains, hydrogen revenue contracts −57.8% due to −67.6% price decline, while $Al(OH)_3$ revenue surges by +233.9% via +98.3% price appreciation, elevating its share from 59.3% to 92.0% as shown in Table 7. Total costs decline −75.4%, dominated by depreciation −80.0% (DDB schedule), with stable OpEx (−0.6% nominal, 3.40 to 3.38 million $). This yields EBIT of $111 million in 2041, total revenue of +114.9%, and a CAGR of 4.3%. Perspective: $Al(OH)_3$ price resilience (+98.3%) compensates H_2 market pressure (−67.6%), with cost deflation (−75.4%) enabling profitability despite early negative cash flows. Sensitivity: ±20% $Al(OH)_3$ price alters IRR by ±4.2 pp.

Market developments transform the economic sequence of products substantially generated from waste resources. A flexible infrastructure should be considered essential, as it enables adjustments to product priorities as market realities evolve. Stronger backing for $Al(OH)_3$ valorisation is received when the government views this process as its central incentive rather than a supplemental value. Through its aluminium slag hydrolysis facility, aluminium companies can demonstrate how technological integrations between AI prediction tools and waste conversion yield strategic resources from what were previously environmental burdens, as utilised in this work. The actual value of $Al(OH)_3$ reuse lies in its ability to develop a system for sustainable resource recovery that can guide national strategies for waste management alongside green energy transitions and industrial expansions.

Research Limitations

The economic analysis predicts that hydrogen costs will decrease by 67.6% while Al $(OH)_3$ prices will increase by 98.3% throughout the 17-year operation. The price projections rely on stable macroeconomic and policy conditions, as hydrogen costs track the global commoditisation of renewable energy. Al $(OH)_3$ prices increase due to the growth of the industrial and other sectors. These assumptions may fail to account for potentially significant risks from geopolitical disturbances, competitive material substitutions, and unanticipated regulatory adjustments. The project's IRR decreases from 13.4% to 9.2% when Al $(OH)_3$ prices fall by 30%, along with an extended payback period stretching from 10 to 14 years, demonstrating the delicate nature of revenue plans based on single-product price growth. The rapid decrease in hydrogen prices will negatively impact project financial returns.

Temporal scope of data and model generalisability. The ML framework applied a Bayesian optimisation framework in Optuna, ran 10,000 iterations, and used early stopping to prevent overfitting, using data from 2005 to 2023. The regularisation features of XGBoost (L1/L2 penalties, subsampling) reduce overfitting, yet the short time span of data makes it difficult for the model to differentiate between permanent trends and short-lived fluctuations. Short-term demand changes, together with policies, can create lasting patterns that the model incorrectly identifies as permanent trends.

The role of experimental validation in enhancing predictive precision. The predictive capabilities of the presented XGBoost model for aluminium production and slag generation remain strong, but direct experimental validation is limited as reaction conditions are yet to be tested. The model relies on stoichiometric conversions combined with empirical ratios, establishing relationships from the literature that may not fully capture the process dynamics that affect hydrogen and $Al(OH)_3$ production at the micro-level, including temperature variations, heterogeneous slag formation, and local reaction rate variations. Experimental

trials with controlled reactors should measure micro-level parameters to yield critical adjustment factors to correct real-world reactivity and yield efficiency deviations. By including specific reaction performance metrics in the ML pipeline, the studied XGBoost model would develop into a stronger and more general model for better forecasting of hydrogen output and material recovery under industrial conditions. Upcoming research efforts need to merge predictive analytics with experimental-empirical studies to refine the algorithm's baseline assumptions and make its output more realistic to technical implementation conditions.

Scalability, slag heterogeneity, and environmental impacts. Although the conceptual process can be scaled, scaling stoichiometric potential to industrial practice introduces hazards, including reactor throughput, mixing and heat removal, gas-liquid separation, and the safe handling of hydrogen at high production rates. The exothermic nature of the NaOH-assisted aluminium hydrolysis system highlights the importance of system-level thermal management to prevent localised overheating, ensure controllable kinetics, and maintain stable operation **[50]**. Previous engineering experience with this system underscores the necessity of practical integration at scale. Additionally, the composition of aluminium slag or dross varies between batches and manufacturers (metallic Al fraction, oxides/spinels, salts, and nitrides/carbides), which can affect conversion efficiency, reaction rate, and the purity or marketability of recovered products. Yield estimates should be considered as upper bounds or scenario-dependent unless site-specific characterisation and pilot trials are performed, in line with the known variability and hazard profile of aluminium dross **[42]**. Any potential environmental impacts related to chemical use and residue management should be addressed directly: chemical alkaline activation may require effluent management for process losses or purge streams, and some dross materials, particularly aluminium nitride (AlN), may hydrolyse and emit ammonia, affecting wastewater treatment and residue handling. Consequently, the proper treatment of process liquids, high-salinity streams, and inert residues should be included in a comprehensive mass-balance approach for the characterisation of effluent and residue management **[64]**.

CONCLUSION

This research proposes a nationwide aluminium slag waste management solution for Saudi Arabia by merging automated aluminium production predictions through ML with hydrolysis-based product extraction. The predictive accuracy of XGBoost models enables precise estimates of aluminium production and slag output to support data-driven infrastructure development, resource allocation, and the design of industrial strategies. The production process uses low-emission, cost-effective methods to generate green hydrogen while producing $Al(OH)_3$ as a primary product from what was previously an industrial byproduct. Research indicates that aluminium waste can be diverted from environmental hazards into industrial applications, thereby establishing sustainable secondary material markets that cut down raw material needs. The solution aligns with Vision 2030 guidelines by developing energy-efficient systems, advancing materials, and advancing environmental sustainability nationwide, while offering a direct implementation approach for transforming waste into valuable assets.

ACKNOWLEDGMENT

The authors are grateful for the support of the Mohammad Al-Aqeel Fellowship Program for Graduate Students Conference Attendance at the King Fahd University of Petroleum and Minerals (KFUPM).

NOMENCLATURE

Greek letters

α	L1 Regularisation Term
λ	L2 Regularisation Term
η	Learning Rate

Subscripts and Superscripts

reg	regularisation

Abbreviations

CAGR	Compound Annual Growth Rate
CapEx	Capital Expenditure
CSTR	Continuous Stirred Tank Reactor
DSS	Decision Support System
DDM	Dividend Discount Model
EBIT	Earnings Before Interest and Taxes
EBITDA	Earnings Before Interest, Taxes, Depreciation, and Amortisation
GDP	Gross Domestic Product
H_2	Molecular hydrogen
IRR	Internal Rate of Return
k-NN	k Nearest Neighbours
ktpa	kilotonnes per annum
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MIRR	Modified Internal Rate of Return
ML	Machine Learning
MSE	Mean Squared Error
Mtpa	Million tonnes per annum
OpEx	Operational Expenditure
R^2	Coefficient of Determination
SCADA	Supervisory Control and Data Acquisition
VIF	Variance Inflation Factor
XGBoost	eXtreme Gradient Boosting

REFERENCES

1. M. Milani and L. Montorsi, Energy Recovery of the Biomass from Livestock Farms in Italy: The Case of Modena Province, *Journal of Sustainable Development of Energy, Water and Environment Systems*, Vol. 6, No. 3, pp 464–480, Sep. 2018, https://doi.org/10.13044/j.sdewes.d6.0199.
2. Q. Zhang, S. Wang, H. Sun, S. G. Arhin, Z. Yang, G. Liu, Y. W. Tong, H. Tian, and W. Wang, Anaerobic digestion + pyrolysis integrated system for food waste treatment achieving both environmental and economic benefits, *Energy*, Vol. 288, 129856, Feb. 2024, https://doi.org/10.1016/J.ENERGY.2023.129856.
3. X. Peng, Y. Jiang, Z. Chen, A. I. Osman, M. Farghali, D. W. Rooney, and P.-S. Yap, Recycling municipal, agricultural and industrial waste into energy, fertilizers, food and construction materials, and economic feasibility: a review, *Environ. Chem. Lett.*, Vol. 21, No. 2, pp 765–801, 2023, https://doi.org/10.1007/s10311-022-01551-5.
4. M. A. Abdelkareem, M. Ayoub, R. I. Al Najada, A. H. Alami, and A. G. Olabi, Hydrogen from waste metals: Recent progress, production techniques, purification, challenges, and applications, *Sustainable Horizons*, Vol. 9, 100079, Mar. 2024, https://doi.org/10.1016/J.HORIZ.2023.100079.

5. D. Stevanovic, C. Kutter, T. Philippi, F. Völkl, and A. Buradkar, Hydrogen Production by Steam Reforming using Biomass, *Journal of Sustainable Development of Energy, Water and Environment Systems*, Vol. 12, No. 2, pp 1–14, Jun. 2024, https://doi.org/10.13044/j.sdewes.d12.0496.
6. K. Lan and Y. Yao, Feasibility of gasifying mixed plastic waste for hydrogen production and carbon capture and storage, *Commun. Earth Environ.*, Vol. 3, No. 1, 300, 2022, https://doi.org/10.1038/s43247-022-00632-1.
7. S. C. Wijayasekera, K. Hewage, O. Siddiqui, P. Hettiaratchi, and R. Sadiq, Waste-to-hydrogen technologies: A critical review of techno-economic and socio-environmental sustainability, *Int. J. Hydrogen Energy*, Vol. 47, No. 9, pp 5842–5870, Jan. 2022, https://doi.org/10.1016/J.IJHYDENE.2021.11.226.
8. M. Aziz, A. Darmawan, and F. B. Juangsa, Hydrogen production from biomasses and wastes: A technological review, *Int. J. Hydrogen Energy*, Vol. 46, No. 68, pp 33756–33781, Oct. 2021, https://doi.org/10.1016/J.IJHYDENE.2021.07.189.
9. L. Ouyang, M. Liu, K. Chen, J. Liu, H. Wang, M. Zhu, and V. Yartys, Recent progress on hydrogen generation from the hydrolysis of light metals and hydrides, *J. Alloys Compd.*, Vol. 910, 164831, 2022, https://doi.org/10.1016/J.JALLCOM.2022.164831.
10. H. Almohamadi, A. L. Khan, A. AlKassem, W. Sindi, S. Alrashdi, and T. Alhazmi, Innovative recycling and conversion of aluminum waste to hydrogen and aluminum chloride: Enhancing economic feasibility and sustainability in Saudi Arabia, *Chemical Engineering Research and Design*, Vol. 212, pp 143–157, Dec. 2024, https://doi.org/10.1016/J.CHERD.2024.10.020.
11. J. Lu, W. Yu, S. Tan, L. Wang, X. Yang, and J. Liu, Controlled hydrogen generation using interaction of artificial seawater with aluminum plates activated by liquid Ga–In alloy, *RSC Adv.*, Vol. 7, No. 49, pp 30839–30844, 2017, https://doi.org/10.1039/C7RA01839H.
12. J. Guo, Z. Su, J. Tian, J. Deng, T. Fu, and Y. Liu, Enhanced hydrogen generation from Al-water reaction mediated by metal salts, *Int. J. Hydrogen Energy*, Vol. 46, No. 5, pp 3453–3463, Jan. 2021, https://doi.org/10.1016/J.IJHYDENE.2020.10.220.
13. M. Su, H. Wang, H. Xu, F. Chen, H. Hu, and J. Gan, Enhanced hydrogen production properties of a novel aluminum-based composite for instant on-site hydrogen supply at low temperature, *Int. J. Hydrogen Energy*, Vol. 47, No. 17, pp 9969–9985, Feb. 2022, https://doi.org/10.1016/J.IJHYDENE.2022.01.092.
14. A. O. Dudoladov, O. A. Buryakovskaya, M. S. Vlaskin, A. Z. Zhuk, and E. I. Shkolnikov, Generation of hydrogen by aluminium oxidation in aquaeous solutions at low temperatures, *Int. J. Hydrogen Energy*, Vol. 41, No. 4, pp 2230–2237, Jan. 2016, https://doi.org/10.1016/J.IJHYDENE.2015.11.122.
15. O. A. Buryakovskaya, M. S. Vlaskin, and S. S. Ryzhkova, Hydrogen production properties of magnesium and magnesium-based materials at low temperatures in reaction with aqueous solutions, *J. Alloys Compd.*, Vol. 785, pp 136–145, May 2019, https://doi.org/10.1016/J.JALLCOM.2019.01.003.
16. D. M. Taher, W. Balmant, S. H. Och, D. B. Pitz, G. S. Venter, L. C. Filho, L. S. Martins, J. V. C. Vargas, and J. C. Ordonez, All-Electric Ship On-Board Continuous Sustainable H2 Generation from Aluminum Scrap and Seawater, in *2023 IEEE Electric Ship Technologies Symposium, ESTS 2023*, Jan. 2023, pp 211–217, https://doi.org/10.1109/ESTS56571.2023.10220471.
17. A. Meng, Y. Sun, W. Cheng, Z. Zhai, L. Jiang, Z. Chong, Y. Chen, and A. Wu, Mechanism of hydrogen generation from low melting point elements (Ga, In, Sn) on aluminum alloy hydrolysis, *Int. J. Hydrogen Energy*, Vol. 47, No. 93, pp 39364–39375, Dec. 2022, https://doi.org/10.1016/J.IJHYDENE.2022.09.127.

18. K. K. Singh, A. Meshram, D. Gautam, and A. Jain, Hydrogen production using waste aluminium dross: from industrial waste to next-generation fuel, *Agronomy Research*, Vol. 17, No. S1, pp 1199–1206, 2019, https://doi.org/10.15159/AR.19.022.
19. B. Das, P. S. Robi, and P. Mahanta, Experimental Investigation and Modelling by Machine Learning Techniques for Hydrogen Generation by Reacting Aluminium with Aqueous NaOH Solution, *Fuel*, Vol. 351, Nov. 2023, https://doi.org/10.1016/j.fuel.2023.128924.
20. E. David and J. Kopac, Hydrolysis of aluminum dross material to achieve zero hazardous waste, *J. Hazard. Mater.*, Vol. 209–210, No. 8, pp 501–509, Mar. 2012, https://doi.org/10.1016/j.jhazmat.2012.01.064.
21. A. Bolt, I. Dincer, and M. Agelin-Chaab, A Review of Unique Aluminum–Water Based Hydrogen Production Options, *Energy & Fuels*, Vol. 35, No. 2, pp 1024–1040, Jan. 2021, https://doi.org/10.1021/acs.energyfuels.0c03674.
22. Z. Hou, X. Cui, and Q. Shi, Prediction of Steel Production Based on the Combination of XGBOOST and LassoLars, *International Conference on Data Science and Information Technology*, pp 143–147, 2021, https://doi.org/10.1145/3478905.3478934.
23. C. Liu, F. Li, P. Zhang, and P. Balasubramanian, Augmented machine learning with limited data for hydrogen yield prediction in wastewater dark fermentation, *npj Clean Water 2025 8:1*, Vol. 8, No. 1, 101, Nov. 2025, https://doi.org/10.1038/s41545-025-00529-4.
24. T. Chen and C. Guestrin, XGBoost: A Scalable Tree Boosting System, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp 785–794, https://doi.org/10.1145/2939672.2939785.
25. A. Al-Ataby, B. N. Getu, H. Attia, and A. Khaimah, Machine Learning-Based Water Quality Prediction, *J.sustain. dev. energy water environ. syst*, Vol. 14, No. 1, 1130634, 2026, https://doi.org/10.13044/j.sdewes.d13.0634.
26. A. T. Hoang, W.-H. Chen, M. O. Guerrero-Pérez, E.-R. Castellón, M. C. López-Escalante, V. N. Nguyen, P. Paramasivam, X. P. Nguyen, and T. H. Truong, Turning waste into energy: Application of machine learning and explainable artificial intelligence for determining key factors in wastewater-to-hydrogen conversion, *Energy*, Vol. 344, 139934, Feb. 2026, https://doi.org/10.1016/j.energy.2026.139934.
27. A. Raghuvira Pratap, S. Panda, and G. Syama Sameera, Predictive Maintenance for Two-Wheeler Vehicles Using XGBoost, in *10th International Conference on Advanced Computing and Communication Systems, ICACCS 2024*, 2024, pp 746–751, https://doi.org/10.1109/ICACCS60874.2024.10717187.
28. L. M. I. Leo Joseph, E. H. Reddy, M. Srinidhi, P. V. Saketh, D. Mohan, and C. Rajendra Prasad, Predicting Real-Time House Prices: A Machine Learning Approach Using XGBoost Algorithm, 2024, https://doi.org/10.1109/APCIT62007.2024.10673571.
29. Q. Yang, L. Zhao, J. Xiao, R. Wen, F. Zhang, and D. Zhang, Machine learning-assisted prediction and optimization of solid oxide electrolysis cell for green hydrogen production, *Green Chemical Engineering*, Vol. 6, No. 2, pp 154–168, Jun. 2025, https://doi.org/10.1016/J.GCE.2024.04.004.
30. Z. Liu, Z. Cui, M. Wang, B. Liu, and W. Tian, A machine learning proxy based multi-objective optimization method for low-carbon hydrogen production, *J. Clean. Prod.*, Vol. 445, 141377, Mar. 2024, https://doi.org/10.1016/J.JCLEPRO.2024.141377.
31. S. Tasneem, A. A. Ageeli, W. M. Alamier, N. Hasan, and M. R. Safaei, Organic catalysts for hydrogen production from noodle wastewater: Machine learning and deep learning-based analysis, *Int. J. Hydrogen Energy*, Vol. 52, pp 599–616, Jan. 2024, https://doi.org/10.1016/J.IJHYDENE.2023.07.114.

32. D. Dong, F. Wen, Y. Zhang, and W. Qiu, Application of XGboost in Electricity Consumption Prediction, in *2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information, ICETCI 2023*, 2023, pp 1260–1264, https://doi.org/10.1109/ICETCI57876.2023.10176934.
33. A. E. A. A. Mohammed *et al.*, Forecasting Mining Revenues in Saudi Arabia: A SARIMA Approach for Economic Diversification under Vision 2030, *Pakistan Journal of Life and Social Sciences (PJLSS)*, Vol. 22, No. 2, 2024, https://doi.org/10.57239/pjlss-2024-22.2.001538.
34. H. Oukhouya, H. Kadiri, K. El Himdi, and R. Guerbaz, Forecasting International Stock Market Trends: XGBoost, LSTM, LSTM-XGBoost, and Backtesting XGBoost Models, *Statistics, Optimization & Information Computing*, Vol. 12, No. 1, pp 200–209, 2024, https://doi.org/10.19139/soic-2310-5070-1822.
35. M. Kahia, T. Moulahi, S. Mahfoudhi, S. Boubaker, and A. Omri, A machine learning process for examining the linkage among disaggregated energy consumption, economic growth, and environmental degradation, *Resources Policy*, Vol. 79, 103104, Dec. 2022, https://doi.org/10.1016/j.resourpol.2022.103104.
36. B. Kenzhaliyev, T. Imankulov, A. Mukhanbet, S. Kvyatkovskiy, M. Dyussebekova, and N. Tasmurzayev, Intelligent System for Reducing Waste and Enhancing Efficiency in Copper Production Using Machine Learning, *Metals (Basel).*, Vol. 15, No. 2, 186, Feb. 2025, https://doi.org/10.3390/met15020186.
37. M. Ran, J. Yi, W. Liu, C. Wang, and Y. Deng, Reinforcement Learning-Based Prediction Method for Aluminum Electrolysis Production, in *2025 37th Chinese Control and Decision Conference (CCDC)*, May 2025, pp 1016–1021, https://doi.org/10.1109/CCDC65474.2025.11090328.
38. Y. Ning, H. Kazemi, and P. Tahmasebi, A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet, *Comput. Geosci.*, Vol. 164, 105126, 2022, https://doi.org/10.1016/j.cageo.2022.105126.
39. M. O. Esangbedo, B. O. Taiwo, H. H. Abbas, S. Hosseini, M. Sazid, and Y. Fissha, Enhancing the exploitation of natural resources for green energy: An application of LSTM-based meta-model for aluminum prices forecasting, *Resources Policy*, Vol. 92, 105014, May 2024, https://doi.org/10.1016/j.resourpol.2024.105014.
40. D. Q. Gbadago, S. Go, and S. Hwang, A leap forward in chemical process design: Introducing an automated framework for integrated AI and CFD simulations, *Comput. Chem. Eng.*, Vol. 192, Jan. 2025, https://doi.org/10.1016/j.compchemeng.2024.108906.
41. F. Guo, Y. Ren, Y. Zhou, S. Sun, M. Cui, and J. Khim, Machine learning vs. statistical model for prediction modeling and experimental validation: Application in groundwater permeable reactive barrier width design, *J. Hazard. Mater.*, Vol. 469, 133825, May 2024, https://doi.org/10.1016/j.jhazmat.2024.133825.
42. M. Mahinroosta and A. Allahverdi, Hazardous aluminum dross characterization and recycling strategies: A critical review, *J. Environ. Manage.*, Vol. 223, pp 452–468, Oct. 2018, https://doi.org/10.1016/J.JENVMAN.2018.06.068.
43. Saudi Arabian Mining Company (Ma'aden), 2023 Annual Report, Riyadh, Saudi Arabia, Dec. 2023.
44. J. P. Hong, J. Wang, H. Y. Chen, B. De Sun, J. J. Li, and C. Chen, Process of aluminum dross recycling and life cycle assessment for Al-Si alloys and brown fused alumina, *Transactions of Nonferrous Metals Society of China (English Edition)*, Vol. 20, No. 11, pp 2155–2161, Nov. 2010, https://doi.org/10.1016/S1003-6326(09)60435-0.
45. A. Meshram and K. K. Singh, Recovery of valuable products from hazardous aluminum dross: A review, *Resour. Conserv. Recycl.*, Vol. 130, pp 95–108, Mar. 2018, https://doi.org/10.1016/J.RESCONREC.2017.11.026.

46. S. K. Verma, V. K. Dwivedi, and S. P. Dwivedi, Utilization of aluminium dross for the development of valuable product – A review, *Mater. Today Proc.*, Vol. 43, pp 547–550, Jan. 2021, https://doi.org/10.1016/J.MATPR.2020.12.045.
47. Z. Zuo, H. Lv, R. Li, F. Liu, and H. Zhao, A new approach to recover the valuable elements in black aluminum dross, *Resour. Conserv. Recycl.*, Vol. 174, 105768, Nov. 2021, https://doi.org/10.1016/J.RESCONREC.2021.105768.
48. P. J. Linstrom and W. G. Mallard, Eds., NIST Chemistry WebBook, SRD 69, 2025. https://doi.org/10.18434/T4D303, [Accessed: Mar. 13, 2025].
49. Y. Ma, A. Preveniou, A. Kladis, and J. B. Pettersen, Circular economy and life cycle assessment of alumina production: Simulation-based comparison of Pedersen and Bayer processes, *J. Clean. Prod.*, Vol. 366, 132807, Sep. 2022, https://doi.org/10.1016/j.jclepro.2022.132807.
50. D. W. Hurtubise, D. A. Klosterman, and A. B. Morgan, Development and demonstration of a deployable apparatus for generating hydrogen from the hydrolysis of aluminum via sodium hydroxide, *Int. J. Hydrogen Energy*, Vol. 43, No. 14, pp 6777–6788, Apr. 2018, https://doi.org/10.1016/J.IJHYDENE.2018.02.087.
51. C. Wang, T. Yang, Y. Liu, J. Ruan, S. Yang, and X. Liu, Hydrogen generation by the hydrolysis of magnesium–aluminum–iron material in aqueous solutions, *Int. J. Hydrogen Energy*, Vol. 39, No. 21, pp 10843–10852, 2014, https://doi.org/10.1016/J.IJHYDENE.2014.05.047.
52. A. Bolt, I. Dincer, and M. Agelin-Chaab, Experimental study of hydrogen production process with aluminum and water, *Int. J. Hydrogen Energy*, Vol. 45, No. 28, pp 14232–14244, May 2020, https://doi.org/10.1016/J.IJHYDENE.2020.03.160.
53. T. Kirton, F. Saceleanu, M. Salehi Mobarakeh, and M. R. Kholghy, Cogeneration of hydrogen, alumina, and heat from aluminum-water reactions, *Int. J. Hydrogen Energy*, Vol. 68, pp 115–127, May 2024, https://doi.org/10.1016/j.ijhydene.2024.04.038.
54. J. Hu, Forecasting copper price using VAR and the XGBoost model: an experiment with a relatively small dataset, M.Sc. Thesis, Lund University, Lund, Sweden, 2023.
55. L. Meroueh, T. W. Eagar, and D. P. Hart, Effects of Mg and Si Doping on Hydrogen Generation via Reduction of Aluminum Alloys in Water, *ACS Appl. Energy Mater.*, Vol. 3, No. 2, pp 1860–1868, Feb. 2020, https://doi.org/10.1021/acsaem.9b02300.
56. L. Meroueh, L. Neil, T. W. Eagar, and D. P. Hart, Leveraging Grain Size Effects on Hydrogen Generated via Doped Aluminum–Water Reactions Enabled by a Liquid Metal, *ACS Appl. Energy Mater.*, Vol. 4, No. 1, pp 275–285, Jan. 2021, https://doi.org/10.1021/acsaem.0c02175.
57. Procurement Resource, Aluminum hydroxide prices, latest price, graph, index, 2025. https://www.procurementresource.com/resource-center/aluminum-hydroxide-price-trends, [Accessed: May 03, 2025].
58. M. Lambert, A. Barnes, A. Marcu, O. Imbault, A. Bhashyam, M. Tengler, C. Cavallera, and G. Romeo, 2024 State of the European Hydrogen Market Report, Oxford Institute for Energy Studies, UK, Jun. 2024.
59. Worthwill, SMM Aluminum Spot Price and Trend Charts, 2024. https://www.worthwillaluminium.com/aluminum-price/smm, [Accessed: May 03, 2025].
60. N. S. A. Zauzi, M. Z. H. Zakaria, R. Baini, M. R. Rahman, N. Mohamed Sutan, and S. Hamdan, Influence of alkali treatment on the surface area of aluminium dross, *Advances in Materials Science and Engineering*, Vol. 2016, 2016, https://doi.org/10.1155/2016/6306304.
61. A. Navarrete and Z. Yuanrong, The price of green hydrogen: How and why we estimate future production costs, *The International Council on Clean Transportation*, May 20, 2024. https://theicct.org/the-price-of-green-hydrogen-estimate-future-production-costs-may24/, [Accessed: May 04, 2025].

62. Future Markets Inc., The Global Hydrogen Market 2025-2035 , Edinburgh, UK, Mar. 2025, https://www.futuremarketsinc.com/the-global-hydrogen-market-2025-2035/ [Accessed: May 04, 2025].
63. E. Taibi, H. Blanco, R. Miranda, and M. Carmo, Green Hydrogen Cost Reduction: Scaling up Electrolysers to Meet the 1.5^0C Climate Goal, Abu Dhabi, United Arab Emirates, Dec. 2020.
64. H. Lv, H. Zhao, Z. Zuo, R. Li, and F. Liu, A thermodynamic and kinetic study of catalyzed hydrolysis of aluminum nitride in secondary aluminum dross, *Journal of Materials Research and Technology*, Vol. 9, No. 5, pp 9735–9745, Sep. 2020, https://doi.org/10.1016/j.jmrt.2020.06.051.

APPENDIX

Data acquisition and preprocessing. The research relies on a systematic approach to acquire and enhance data from multiple sources to establish reliable foundations for predictive modelling. The research activities focused on discovering multiple elements that affect aluminium production in Saudi Arabia through economic analysis and operational and policy examinations. Ma'aden's annual reports from 2005 to 2023 provided detailed statistics on primary production, including annual aluminium output, efficiency rates and regional operational measurements. National industrial bulletins issued by the Saudi Ministry of Energy provided information on Vision 2030 resource distribution along with production capacity updates. GDP growth figures, along with industrial sector data and energy price statistics, were acquired from Ma'aden's report.

The modelling of autocorrelation and short-term volatility reduction utilised temporal variables that included 1–3 year lagged output values, combined with three-year rolling averages. The analysis multiplied energy prices by industrial GDP growth to explain how economic activity and production costs interact. The 2016 launch of Vision 2030 was treated as a binary variable treatment to detect industrial strategic changes within this policy framework. The model received continuous variables after min-max scaling to eliminate scale bias and processed categorical variables through one-hot encoding to meet model compatibility requirements. The k-nearest neighbours (k=3) imputation method was selected to fix data deficiencies in less than 5% of the dataset because it preserves both temporal and contextual relationships inside the dataset. The alternative imputation methods used for cross-validation confirmed that k-NN produced consistent results since model outputs fluctuated by less than 2%. The final dataset, containing 19 annual observations from 2005 to 2023, included 15 engineering features, and multicollinearity was evaluated using variance inflation factors (VIF < 5). The systematic approach to data collection and preprocessing created an accurate model while ensuring replicability across similar industrial forecasting situations.

XGBoost architecture and hyperparameter optimisation. Min-max scaling normalisation applied to all continuous variables enabled model training and convergence. One-hot encoding transformed the categorical and binary variables, which included project milestones. The structured dataset was organised into rows representing individual years, each containing both input variables and their corresponding aluminium production targets. The training process utilised data points from 2005 to 2018, while validation was performed using data points from 2019 to 2023. The chronological order of data was maintained using a split method that prevented information from crossing the test and train periods.

Table A1. Model parameters

Parameter	Search space	Optimal value
n_estimators	100–1000	500
max_depth	3–12	8
learning_rate	0.01–0.3	0.05
subsample	0.6–1.0	0.8
colsample_bytree	0.5–1.0	0.7
reg_alpha	0–1	0.5
reg_lambda	0–2	1

The XGBoost regressor was implemented in Python using the xgboost library after performing randomised-search optimisation of hyperparameters (**Table A1**). These included maximum tree depth (max_depth), learning rate (η), and number of estimators (n_estimators), and subsampling ratio (subsample) and column sampling ratio (colsample_bytree). The XGBoost regressor uses a framework that optimises model complexity and processing performance. The foundational model received specific parameters for aluminium production data characteristics, including a regression error minimisation function (reg:squarederror) and 500 decision trees (n_estimators) with a maximum tree depth (max_depth) limit to achieve accurate predictions. Gradient descent optimisation used a learning rate of 0.05, while subsampling was set to 0.8 and column sampling was set to 0.7 to enhance model generalisation. L1 regularisation with reg_alpha set to 0.5 and L2 regularisation with reg_lambda set to 1.0 were used to remove unimportant features while optimising weight distribution.

A Bayesian optimisation framework in Optuna ran 10,000 iterations to minimise Mean Absolute Percentage Error (MAPE) in hyperparameter optimisation. The search area included parameters that focused on tree depth (max_depth: 3–12), learning rate (learning_rate: 0.01–0.3), and subsampling ratios (subsample: 0.6–1.0). The optimal configuration was max_depth=8, which provided a good compromise between model complexity and interpretability, and learning_rate=0.05, which stabilised error reduction. Column sampling (colsample_bytree=0.7) reduced model overfitting by randomly selecting features for tree construction. The training terminated via early stopping, with patience=10 to monitor the validation loss and stop training when improvements stalled, to prevent overfitting and minimise resource usage.

Model training and regularisation. The XGBoost model training process implemented specific methods to improve predictive accuracy and counteract overfitting, given the dataset's short time span. The gradient-boosting system created an accumulative set of decision trees that focused on fixing errors remaining from previous execution steps. Through sequential learning, the model established advanced relationships between energy prices and annual aluminium production, as well as between historical trends and target data. The model gained generalisability through stochastic gradient descent, with random observation and feature sampling, where each tree received 80% of observations and 70% of features. The selected subsampling strategy introduced controlled variability into the training data, making the model less vulnerable to noisy industrial and economic data.

Model complexity was controlled through the systematic application of regularisation techniques, thereby improving the model's interpretability. The L1 regularisation method, with a penalty coefficient of 0.5, successfully set unimportant features to zero during automatic feature selection. The L2 (Ridge) regularisation method was applied with λ=1.0 to penalise the weights, preventing substantial prediction errors and addressing variable

collinearity issues. Model parsimony and predictive effectiveness existed in equilibrium because dual regularisation worked together as a system.

The optimisation system minimised the Mean Squared Error (MSE) loss function because it works best with outliers and supports gradient-based learning. At each iteration, the loss function computed its partial derivatives to guide the formation of decision trees toward splits that maximised residual error reduction. The learning procedure terminated early via early stopping when the validation loss converged after 10 epochs, helping prevent overfitting and maximising resource efficiency. The applied methodology generated a model that showed strong agreement with the economic and time-based patterns of Saudi Arabia's aluminium industry, with a test-set MAPE of 6.9% and an R^2 of 0.98.

Estimating aluminium slag generation using predictive outputs. The calculation of aluminium slag output relies on proven empirical conversion factors that relate aluminium production to the quantities of byproduct slag. The smelting methods used in Gulf Cooperation Council (GCC) nations produce primary aluminium slag, accounting for 15–25% of total manufacturing mass. The specified range illustrates variations between different smelting methods and raw material quality levels, as well as operational performance standards. The slag yield from high-purity alumina smelters in Saudi Arabia ranges from 18% to 22%, based on metallurgical analysis of slag composition, according to Ma'aden management data. Expert analysis showed that slag primarily contains metallic aluminium (12–18%) and aluminium oxide (50–65%), with minimal nitrides and carbides, in addition to small variations resulting from furnace settings and alloy formation. The analysis established a specific conversion rate of 20% for accurate, actionable predictions that match the highest recorded values reported in Resources, Conservation & Recycling and validate Ma'aden's historical data from 2015 through 2023. The rate utilises real-world findings to combine technical accuracy with practical considerations, managing natural process variations to establish a solid foundation for the development of a waste management strategy.

Paper submitted: 19.12.2025
Paper revised: 15.05.2026
Paper accepted: 19.05.2026