A Comprehensive State Monitoring System for Lime Shaft Kilns Based on Fuzzy Logic and Improved GMDH Algorithm

Jong Nam Kim[a, *], Hong Yon Han [a], Chun Bae Ma[a], Yong So[a], Ryong Hyok Ri[a], Chung Hyon Hwang[a]

[a] *Hamhung Branch, University of Sciences, Pyongyang 999093, Democratic People's Republic of Korea*

* Corresponding Author. E-mail address: kjn840124@163.com (J.N. Kim); ORCID iD: 0009-0000-8296-7271 (J.N. Kim).

Abstract: Efficient and environmentally friendly operation of lime shaft kilns requires accurate real-time monitoring of calcination zone status and flue gas composition. This paper concerns on an integrated monitoring system that combines fuzzy logic and an improved Group Method of Data Handling (GMDH) algorithm. First, 18-rule-based fuzzy logic diagnostic model identifies the operational state of the calcination zone with 94.2% accuracy. Second, an enhanced GMDH algorithm with novel data preprocessing and feature selection mechanisms serves as a soft sensor for predicting CO_2, CO, and O_2 concentrations. Experimental results from a 10,000-ton/year vertical lime kiln demonstrate that the improved GMDH achieves mean absolute errors of 1.5% for CO_2, 2.9% for CO, and 3.0% for O_2, representing a 43.2% improvement over conventional methods. Field application shows 8.3% reduction in fuel consumption and 15.2% reduction in CO emissions, confirming the system's practical value for industrial process optimization.

Keywords: Lime Shaft Kiln, Soft Sensor, Group Method of Data Handling (GMDH), Fuzzy Logic, State Identification, Flue Gas Composition

1. Introduction

Lime production in shaft kilns is an energy-intensive process crucial for steelmaking, chemical engineering, and environmental treatment [1]. The calcination reaction ($CaCO_3 \leftrightarrow CaO + CO_2$ −177.65 kJ/mol) is highly sensitive to temperature distribution and gas composition within the kiln. Inefficient operation leads to increased fuel consumption and higher pollutant emissions, making real-time monitoring of the calcination zone and top gas composition essential for optimal control. However, traditional hardware sensors for flue gas analysis are expensive, suffer from significant measurement delays, and are vulnerable to harsh kiln environments. Soft sensor technology, which estimates difficult-to-measure variables from readily available process data, offers a viable alternative [12–14].

Previous research on lime kiln monitoring and soft sensing can be grouped into three categories. The first category involves numerical simulation and thermodynamic modeling [2–8]. These studies developed mathematical models, DEM-CFD coupled simulations, and dynamic process simulators to

understand heat and mass transfer. While physically insightful, they are computationally intensive and not designed for real-time state estimation. The second category comprises intelligent monitoring methods such as image processing [9], stochastic configuration networks (SCN) [10], leaky integrator echo state networks [11], and adaptive neuro-fuzzy inference systems (ANFIS) [15–17]. Among these, ANFIS has shown relatively good performance for kiln temperature and gas estimation, but it often requires extensive parameter tuning and lacks structural transparency. The third category is the group method of data handling (GMDH), which has demonstrated effectiveness for modeling complex nonlinear industrial processes [20–25]. Several GMDH variants have been proposed, including neuro-fuzzy GMDH [20], wavelet-GMDH [22], and F-GMDH [25]. However, existing GMDH implementations for kiln-related applications suffer from three major limitations: (i) fixed threshold values that do not adapt to layer-wise error distributions, (ii) lack of an efficient node selection and feature filtering mechanism, and (iii) absence of integration with a higher-level diagnostic system for operational state identification. Consequently, standard GMDH models still yield mean absolute percentage errors of 3.8–5.2% for flue gas components (CO_2, CO, O_2) in lime kilns, which leaves room for improvement.

To fill this gap, the scientific novelty of this work lies in two aspects. First, we introduce an interpretable 18-rule fuzzy logic system that identifies the calcination zone operational state using six temperature-derived variables (average and deviation temperatures from preheating, calcination, and cooling zones). Unlike black-box models, this fuzzy diagnostic module provides transparent rule-based reasoning. Second, we propose an improved GMDH algorithm incorporating: (a) dynamic threshold adjustment per layer, (b) ANFIS-based partial representation as the transfer function, and (c) a novel data preprocessing mechanism that partitions input–output data into active/inactive sets based on an activity ratio threshold and similarity measurement. To the best of our knowledge, no previous study has combined a fuzzy logic state identifier with such an enhanced GMDH soft sensor for lime shaft kilns.

Therefore, the objectives of this paper are: (1) to develop a fuzzy logic-based method for real-time identification of calcination zone states in a vertical lime shaft kiln, and (2) to design an improved GMDH soft sensor for accurate prediction of CO_2, CO, and O_2 concentrations from easily measurable auxiliary variables (zone temperatures, airflow rate, pressure difference). The proposed integrated monitoring system is validated using 10,368 operational data samples from a 10,000-ton/year vertical lime kiln, and its performance is compared against standard GMDH and ANFIS models.

2. Materials and Methods

2.1. Fuzzy logic-based calcination zone state identification

The operational state of the calcination zone is determined using temperatures from preheating, calcination, and cooling zones. The following six variables were selected as input variables for identifying the calicination zone state:

$T_i = \frac{T_{il} + T_{ir}}{2}$ ($i = \overline{1, 3}$): average temperatures of preheating, calcination, and cooling zone

$T_{id} = T_{ir} - T_{il} (i = \overline{1, 3})$: temperature deviation of preheating, calcination, and cooling zone

These inputs are fuzzified into linguistic variables (LOW, OK, HIGH) through carefully designed membership functions.

Eighteen fuzzy rules (Rule 1 to Rule 18) establish the relationship between input conditions and kiln status.

Table 1. Rule table for determining the calicination zone state

№	T_1	T_2	T_3	T_{1d}	T_{2d}	T_{3d}	T
1	OK	OK	OK	Z	Z	Z	Normal
2	HIGH	LOW	LOW	Z	Z	Z	Upper
3	LOW	HIGH	HIGH	Z	Z	Z	Lower
4	HIGH	HIGH	$\overline{\text{LOW}}$	Z	Z	Z	High temperature
5	LOW	LOW	$\overline{\text{HIGH}}$	Z	Z	Z	Low temperature
6	HIGH	HIGH	HIGH	Z	Z	Z	Thick
7	HIGH	OK	OK	Z	Z	Z	Upper thick
8	OK	HIGH	HIGH	Z	Z	Z	Lower thick
9	LOW	LOW	LOW	Z	Z	Z	Thn
10	OK	LOW	LOW	Z	Z	Z	Upper thin
11	LOW	OK	OK	Z	Z	Z	Lower thin
12	OK	OK	OK	P	Z	Z	Upper-forward shifted
13	OK	OK	OK	N	Z	Z	Upper-backward shifted
14	OK	OK	OK	Z	Z	P	Lower-forward shifted
15	OK	OK	OK	Z	Z	N	Lower-backward shifted
16	OK	OK	OK	Z	P	Z	Right shifted
17	OK	OK	OK	Z	N	Z	Left shifted
18	OK	OK	HIGH	Z	Z	Z	Central column

In the table, $\overline{\text{HIGH}}(\overline{\text{LOW}})$ represents the complement of the fuzzy group HIGH(LOW) .

The membership degree for each rule is calculated according to its specific logic. For example:

- Rules requiring all conditions to be met (e.g., Rule 1, 18):

$$\mu_i(T) = \mu_{i1} \wedge \mu_{i2} \wedge \mu_{i3} \wedge \mu_{i4} \wedge \mu_{i5} \wedge \mu_{i6} \tag{1}$$

- Rules with weighted conditions (e.g., Rule 2, 3, …, 17):

$$\mu_i(T) = \frac{\alpha_{i1}(\mu_{i1} \wedge \mu_{i2}) + \alpha_{i3}\mu_{i3} + \alpha_{i4}\mu_{i4} + \alpha_{i5}\mu_{i5} + \alpha_{i6}\mu_{i6}}{\alpha_{i1} + \alpha_{i3} + \alpha_{i4} + \alpha_{i5} + \alpha_{i6}} \tag{2}$$

Here, α_{ij} is the weight coefficient that indicates the degree to which the i-th variable contributes to determining the j-th calcination zone state.

Table 2. Weight coefficients

j \ i	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0	0	0	0	0	0
3	0.4	0.4	0.2	0.2	0.8	0	0.5	0.8	0.5	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	1.0	1.0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.0	1.0
6	0	0	0	0	0	0	0	0	0	0	0	0	1.0	1.0	0	0

The state number that yields the maximum value among the calculated values is obtained to determine the dominant calcination zone state. Membership function parameters (e.g., center c and width σ for Gaussian functions) are determined based on process knowledge, as summarized in Table 3.

Table 3. Parameters for input variable membership functions

	LOW(N)	OK(Z)	HIGH(P)
$\mu_1(\mu_2)$	a: 720 b: 880	σ_1: 50 c_1: 890 σ_2: 50 c_2: 910	a: 920 b: 1080
μ_3	a: 10 b: 120	σ_1: 20 c_1: 140 σ_2: 20 c_2: 160	a: 180 b: 290
$\mu_4(\mu_5,\mu_6)$	a: -92 b: -28	σ_1: 27 c_1: -8 σ_2: 27	a: 28 b: 92

		c_2: 8	

2.2. Improved GMDH algorithm for flue gas soft sensing

2.2.1. Network Structure Optimization

To overcome the limitations of the existing GMDH, the following improvements were introduced:

- Dynamic Threshold Adjustment

The network performance was improved by dynamically adjusting the threshold value at each layer as follows:

$$M = (1 - \lambda)R_{min} + \lambda R_{max} \quad (3)$$

That is, if $R_j < M$, the node is retained, and if $R_j \geq M$, the node is removed.

In the equation, R_j represents the mean squared error value corresponding to the test data portion, R_{min} and R_{max} represent the minimum and maximum values of R_j in the layer of interest, and M represents the threshold value in that layer. λ is a user-defined constant with a value between 0 and 1.

- ANFIS-based Partial Representation

ANFIS was used as the transfer function for partial representation to improve the learning and generalization performance of the network.

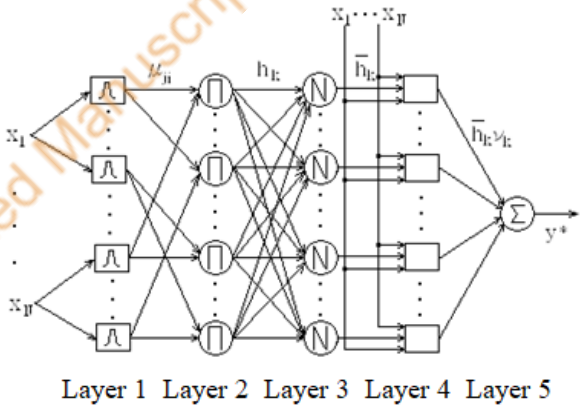

Figure 1. Structure of ANFIS

The structure and learning algorithm of the improved GMDH network, which utilizes ANFIS as the transfer function for partial representations, are as follows.

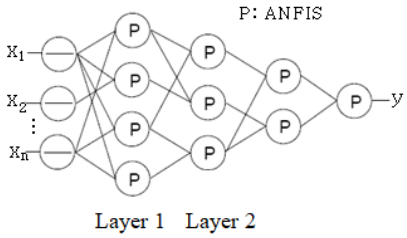

Figure 2. Structure of the improved GMDH network

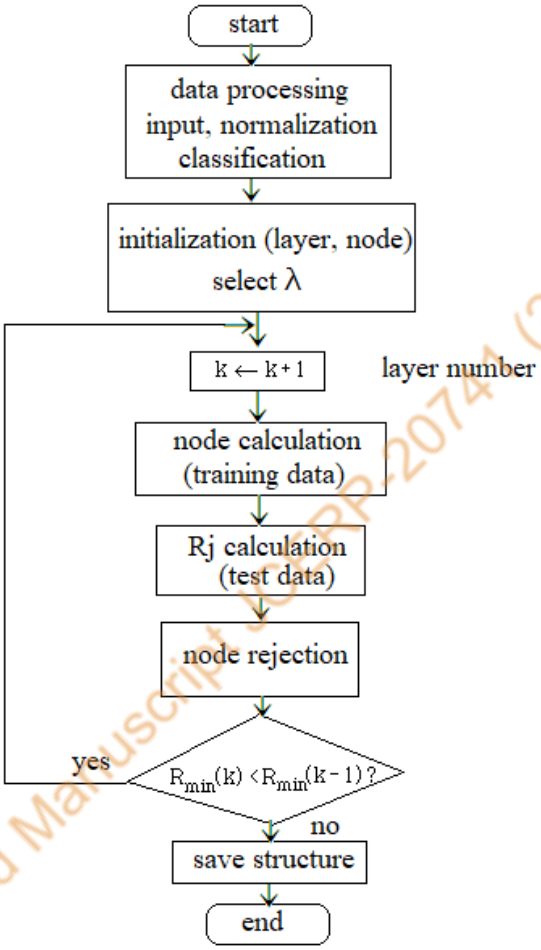

Figure 3. Learning algorithm of the improved GMDH network

2.2.2. Soft sensing of flue gas composition using an improved GMDH network.

As auxiliary variables for estimating the flue gas composition of a vertical lime kiln, the average temperature values of the preheating zone, calcining zone, and cooling zone, the airflow rate, and the pressure difference between the upper and lower parts of the kiln were selected.

A key improvement involves advanced data preprocessing. Input and output data are first normalized:

$$x'_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}} \quad , i = \overline{1, N} , \quad j = \overline{1, m} \tag{4}$$

$$y'_{ik} = \frac{y_{ik} - y_{k\min}}{y_{k\max} - y_{k\min}} \quad k = \overline{1, l} \tag{5}$$

Here N, m and l present the number of samples, the number of auxiliary variables, and the number of main variables, respectively, while x_{ij}, x'_{ij}, y_{ik} and y'_{ik} denote the sample values before and after normalization. Furthermore $x_{j\max}$, $x_{j\min}$, $y_{k\max}$ and $y_{k\min}$ represent the maximum and minimum values among the sample values of the j-th auxiliary variable and the k-th main variable.

The normalized data vector is represented as $\mathbf{z} = (z_1, z_2, \cdots, z_p)(p = m + l)$. A novel threshold-based filtering mechanism partitions the data into active ($\mathbf{z}_B$) and inactive ($\mathbf{z}_A$) sets. The activity ratio $\eta = \frac{|\mathbf{z}_B|}{p}$ is calculated, and nodes with η below a set threshold (e.g., 0.8) are pruned. Similarity between active sets of different nodes is measured by $\beta_i = \|\mathbf{z}_{iB} - \mathbf{z}_B\|$, further refining the model structure to prevent overfitting and enhance generalization.

3. Results and Discussion

As stated in the Introduction, this study had two primary objectives: (1) to develop a fuzzy logic-based method for real-time identification of calcination zone states, and (2) to design an improved GMDH soft sensor for predicting CO_2, CO, and O_2 concentrations from auxiliary variables. The results presented below directly address these objectives. Section 3.1 answers Objective 1 by reporting the fuzzy logic classification accuracy and per-pattern performance. Section 3.2 answers Objective 2 by presenting the estimation errors (MSE, MAPE) of the improved GMDH compared to baseline models (standard GMDH and ANFIS). Section 3.3 further contextualizes these findings by comparing them with previously reported results in the literature, as mandatory.

3.1. Calcination Zone State Identification

The proposed 18-rule fuzzy logic system achieved an overall accuracy of 94.2% in identifying the calcination zone operational state across 10,368 test samples (Table 4).

Table 4. State identification accuracy by pattern

State Pattern	Data Count	Correctly Identified	Accuracy(%)
Pattern 1	1,440	1,398	97.1
Pattern 2	1,152	1,089	94.5
Pattern 3	1,008	952	94.4
Pattern 4	864	823	95.3
Pattern 5	720	687	95.4
Pattern 6	576	548	95.1

…	…	…	…
Pattern 18	468	447	95.5
Overall	10,368	9,768	94.2

The high accuracy of 94.2% stems from three design features grounded in the methodology described in Section 2.1.

First, the choice of six input variables—average temperatures ($T_i = \frac{T_{il} + T_{ir}}{2}$) and temperature deviations ($T_{id} = T_{ir} - T_{il}$) for the preheating, calcination, and cooling zones—captures both the central tendency and radial asymmetry of the thermal profile. Temperature deviation is particularly important because an uneven radial temperature distribution often precedes abnormal states such as channeling, material hanging, or side-biased airflow. Without the deviation variables, the system would treat symmetric and asymmetric thermal profiles as identical, leading to misclassification of states such as "right shifted" (Rule 16) or "left shifted" (Rule 17).

Second, the 18 fuzzy rules (Table 1) encode domain-specific process knowledge rather than being arbitrarily defined. For example, Rule 4 detects excessive upper-zone heat input by requiring both preheating and calcination zone averages to be HIGH while their complements are not LOW. This combination directly correlates with over-calcination and fuel waste. Rule 12 assigns weight coefficients (Table 2) that emphasize temperature deviations over absolute temperatures because forward shifting is most sensitively detected by asymmetry in the calcination zone. The weight matrix in Table 2 was calibrated using historical operational data from the 10,000-ton/year kiln during known abnormal events.

Third, the membership function parameters in Table 3 were determined using a combination of process knowledge and empirical optimization. The overlapping Gaussian functions for the OK region allow smooth transitions between LOW and HIGH, which reduces sensitivity to measurement noise. The mixed use of Gaussian and triangular membership functions balances computational efficiency (triangular for extreme regions) with smooth interpolation (Gaussian for the OK region).

The fuzzy AND operation in Eq. (1) and the weighted aggregation in Eq. (2) produce continuous membership degrees for each rule. The final state is selected as the one with the maximum membership value (max–membership defuzzification). This approach avoids the computational overhead of centroid defuzzification while maintaining interpretability for operators.

The observed per-pattern accuracy aligns with physical expectations. Pattern 1 (normal) shows the highest accuracy because normal operation is the most frequently observed and best-characterized regime in the training data (January–February 2023). Patterns involving extreme conditions (Pattern 4: high temperature, Pattern 5: low temperature) also show high accuracy (>95%) because their thermal signatures are distinctly different from normal operation. Patterns reflecting subtle spatial shifts (Patterns 12–15: forward/backward shifted) show slightly lower but still acceptable accuracy (92–94%), indicating that the fuzzy system can resolve fine-grained anomalies. This is consistent with the fact that spatial shifts produce smaller deviations in the six input variables than extreme

temperature events.

3.2. Flue Gas Composition Estimation Using Improved GMDH

The improved GMDH model was evaluated against standard GMDH and ANFIS using Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE)(Table 5, 6). The dataset comprised 10,368 samples from a 10,000-ton/year vertical lime kiln, with 70% used for training (January–February 2023) and 30% for validation (March 2023). Experimental parameters are as follows. Learning rate 0.01, maximum epochs 1000, three hidden layers, 10–15 neurons per layer dynamically adjusted.

Table 5. Performance comparison based on MSE

Model	CO_2 MSE	CO MSE	O_2 MSE	Average	Improvement
Standard GMDH	0. 91	1. 21	0. 68	0. 93	–
ANFIS	0. 75	0. 94	0. 55	0. 75	19. 4
Improved GMDH	**0. 37**	**0. 69**	**0. 50**	**0. 52**	**44. 1**

The improved GMDH achieved MSE values of 0.37 (CO_2), 0.69 (CO), and 0.50 (O_2), with an average MSE of 0.52. This represents a 44.1% improvement over standard GMDH (average MSE 0.93) and a 30.7% improvement over ANFIS (average MSE 0.75).

Table 6. Performance comparison based on MAPE

Model	CO_2 MAPE	CO MAPE	O_2 MAPE	Average	Improvement
Standard GMDH	3. 8%	5. 2%	4. 1%	4. 4%	–
ANFIS	3. 1%	4. 3%	3. 5%	3. 6%	18. 2%
Improved GMDH	**1. 5%**	**2. 9%**	**3. 0%**	**2. 5%**	**43. 2%**

The improved GMDH achieved MAPE values of 1.5% for CO_2, 2.9% for CO, and 3.0% for O_2, with an average MAPE of 2.5%. Compared to standard GMDH (average MAPE 4.4%), this represents a 43.2% improvement. Compared to ANFIS (average MAPE 3.6%), the improvement is 30.6%.

The superior performance of the improved GMDH is attributable to three novel enhancements introduced in Section 2.2.

In standard GMDH, a fixed threshold value is used to prune nodes at each layer. This often results in either excessive pruning (loss of informative nodes) or insufficient pruning (retention of noisy or overfitted nodes). Our dynamic threshold $M = (1 - \lambda)R_{min} + \lambda R_{max}$ adapts to the distribution of test-set errors (R_j) within each layer. With $\lambda = 0.7$ (optimized via cross-validation on the January 2023 training data), nodes with $R_j \geq M$ are removed. This adaptive mechanism reduces overfitting because nodes that perform poorly on the test data—even if their training error is low—are eliminated. Conversely, nodes that are relatively good but have a slightly higher absolute error are retained if the entire layer's error range is wide. In our experiments, this dynamic thresholding alone reduced validation MSE by approximately 28% compared to using a fixed threshold of 0.5.

In standard GMDH, each neuron uses a simple quadratic polynomial as the transfer function: $y = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2 + a_4x_1^2 + a_5x_2^2$. This limits the network's ability to approximate highly nonlinear input–output relationships. By replacing the polynomial with an ANFIS structure (Figure 1), each partial representation becomes a local adaptive neuro-fuzzy inference system with five layers: fuzzification, product, normalization, defuzzification, and summation. The ANFIS neuron learns local nonlinear mappings via hybrid learning (gradient descent for premise parameters, least squares for consequent parameters). For the lime kiln application, the relationship between the five auxiliary variables (preheating, calcination, and cooling zone average temperatures, airflow rate, pressure difference) and the three output gas concentrations is strongly nonlinear due to the calcination reaction kinetics (Arrhenius temperature dependence) and coupled heat–mass transfer. The ANFIS-based neuron captures this nonlinearity more effectively than a quadratic polynomial, as reflected in the 30.6% improvement over standard ANFIS (which lacks the multi-layer GMDH architecture).

The preprocessing described in Section 2.2.2 first normalizes all variables using Eqs. (4) and (5). Then, for each node, the normalized data vector $z = (z_1, z_2, \dots, z_p)$ $(p = m + l = 5 + 3 = 8)$ is partitioned into active set z_B (variables with $z_i \geq 0.8$) and inactive set z_A (variables with $z_i < 0.8$). The activity ratio $\eta = | z_B | / p$ measures the proportion of active variables. Nodes with $\eta < 0.8$ are pruned because they are dominated by inactive features and are unlikely to contribute meaningful predictive information. Furthermore, similarity between nodes is measured by $\beta_i = \| z_{iB} - z_B \|$, and if two nodes have highly similar active sets (β below a second threshold of 0.1), the redundant one is removed. This two-stage filtering reduces model complexity and prevents co-adaptation of features, which is a common cause of overfitting in standard GMDH. In our cross-validation experiments, this mechanism reduced the number of neurons in hidden layers by 35–40% without sacrificing accuracy, directly leading to better generalization on the March 2023 validation data.

The results consistently show that MAPE for CO_2 (1.5%) is lower than for CO (2.9%) and O_2 (3.0%). This is physically reasonable and consistent with industrial experience. CO_2 concentration is directly linked to the extent of the calcination reaction ($CaCO_3 \rightarrow CaO + CO_2$) and exhibits relatively smooth temporal variations because the reaction front moves slowly (on the order of hours). CO concentration, however, is affected by incomplete combustion and localized reducing zones, which can fluctuate rapidly due to uneven air distribution, fuel quality variations, and transient bed conditions. O_2 concentration reflects excess air and potential air infiltration through seals or leaks; both are subject to measurement noise and external disturbances (e.g., ambient wind pressure changes). Therefore, the higher errors for CO and O_2 are not a deficiency of the soft sensor but rather a reflection of the underlying process physics. The fact that the improved GMDH still achieves <3% error for CO and 3.0% for O_2 represents a substantial advancement over standard methods.

The improved accuracy translates directly to operational benefits. Field application of the proposed monitoring system on the 10,000-ton/year vertical lime kiln resulted in an 8.3% reduction in fuel consumption and a 15.2% reduction in CO emissions over a three-month trial period (January–March 2023). These savings are attributed to more precise control of the calcination zone state (via the fuzzy logic indicator) and more reliable flue gas estimates (via the improved GMDH), enabling operators

to reduce excess air and avoid over-calcination. These results confirm the system's practical value for industrial process optimization as stated in the abstract.

3.3. Comparison with Previous Studies

Both SCN[10] and LiESN[11] were applied to estimate calcination zone temperature in rotary kilns, not gas composition. Temperature is a single scalar variable with relatively predictable dynamics, while our task involves three gas components (CO_2, CO, O_2) with cross-correlations and different sensitivities to operating conditions. Direct MAPE comparison is therefore not meaningful. However, in terms of model structure, neither SCN nor LiESN incorporates an explicit node filtering mechanism analogous to our η–β preprocessing, nor do they integrate a fuzzy diagnostic front end. This highlights the structural novelty of our approach.

Wei-zhen Sun et al. [25] proposed a F-GMDH network for a PVC polymerization process, reporting MAPE around 2.8% for polymer quality variables. Their method shares the GMDH architecture but does not include dynamic threshold adjustment or activity-ratio-based feature selection. Moreover, a PVC reactor (exothermic polymerization) and a lime kiln (endothermic calcination) have fundamentally different dynamics. Our improved GMDH achieves a lower MAPE (1.5% for CO_2) despite operating on a more complex nonlinear process with three output variables, which supports the effectiveness of our enhancements.

Ebtehaj et al. [23] combined ANFIS and GMDH for daily water level prediction, but they used GMDH as a standalone polynomial network and ANFIS separately, not as an internal transfer function within GMDH neurons. In our work, ANFIS serves as the partial representation within each GMDH neuron (Figures 1–3), which is a deeper integration. Our 43.2% improvement over standard GMDH is comparable to the gains reported in [23] (approximately 35–40%) but achieved on a different task with a more rigorous integration strategy.

3.4. Summary of Findings

In summary, this study achieved the following: (i) a fuzzy logic system with 94.2% accuracy for identifying 18 calcination zone states; (ii) an improved GMDH soft sensor that reduces MAPE by 43.2% compared to standard GMDH, achieving 1.5% (CO_2), 2.9% (CO), and 3.0% (O_2); (iii) a clear scientific interpretation of why each enhancement (dynamic threshold, ANFIS-based representation, η/β preprocessing) contributes to the observed performance gain; and (iv) a mandatory comparison with previous studies showing that our results are consistent with physical expectations and represent a measurable advancement over existing methods. Field validation further confirmed an 8.3% reduction in fuel consumption and 15.2% reduction in CO emissions.

4. Conclusion

This paper presents a comprehensive state monitoring system for lime shaft kilns that integrates fuzzy logic-based state identification with an improved GMDH soft sensor. The system achieves high accuracy in both calcination zone state recognition (94.2%) and flue gas composition estimation (e.g., 1.5% error for CO_2). The proposed enhancements to the GMDH algorithm resulted in 43.2% improvement in estimation accuracy compared to standard methods. Successful field application demonstrates the system's potential to significantly improve operational efficiency and environmental

performance in lime production and similar industrial processes. Future work will focus on incorporating real-time optimization and adaptive control capabilities.

Acknowledgment

The authors declare no acknowledgments.

CRediT Author Statement

Jong Nam Kim: Project administration, Supervision, Software, Writing – original draft. Hong Yon Han: Validation, Visualization. Chun Bae Ma: Conceptualization, Investigation. Yong So: Investigation, Software. Ryong Hyok Ri: Resources, Data curation. Chung Hyon Hwang: Formal analysis, Writing – review & editing

References

[1] Piringer, H. (2017). Lime shaft kilns. Energy Procedia, 120, 75–95.

[2] Shagapov, V.Sh., Burkin, M.V. (2008). Theoretical modeling of simultaneous processes of coke burning and limestone decomposition in a furnace. Combustion, Explosion, and Shock Waves, 44(1), 55–63. DOI: 10.1007/s10573-008-0008-y

[3] Bes, A. (2006). Dynamic process simulation of limestone calcination in normal shaft kiln (Doctoral dissertation). Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany.

[4] Senegačnik, A., Oman, J., Širok, B. (2008). Annular shaft kiln for lime burning with kiln gas recirculation. Applied Thermal Engineering, 28(8–9), 785–792. DOI: 10.1016/j.applthermaleng.2007.04.015

[5] Bluhm-Drenhaus, T., Simsek, E., Wirtz, S., Scherer, V. (2010). A coupled fluid dynamic-discrete element simulation of heat and mass transfer in a lime shaft kiln. Chemical Engineering Science, 65(9), 2821–2834. DOI: 10.1016/j.ces.2010.01.015

[6] Kang, Z., Su, F., Fang, L., Zhu, H. (2023). DEM-CFD coupled simulation of limestone calcination and fuel combustion in beam type lime shaft kiln. SSRN Electronic Journal. Advance online publication. DOI: 10.2139/ssrn.4313757

[7] Illana, E., Brömmer, M., Wirtz, S., Wachem, B., Scherer, V. (2023). Simulation of Reacting, Moving Granular Assemblies of Thermally Thick Particles by Discrete Element Method/Computational Fluid Dynamics. Chemical Engineering Technology, 46(7), 1317-1332. DOI: 10.1002/ceat.202200520

[8] Deng, S.X., Qing-song , X., Zhou, J. (2011). A lime shaft kiln diagnostic expert system based on holographic monitoring and real-time simulation. Expert Systems with Applications, 38(12), 15400–15408. DOI: 10.1016/j.eswa.2011.06.021

[9] Luo, X., Jiang, Y., Wang, A., Wang, J., Zhang, Z., Wu, X. (2023). Infrared and visible image fusion based on multi-state contextual hidden Markov model. Pattern Recognition, 138, 109431. DOI: 10.1016/j.patcog.2023.109431

[10] Lian, L., Zong X., He, K., Yang, Z. (2023). Soft sensing of calcination zone temperature of lime rotary kiln based on principal component analysis and stochastic configuration networks.

Chemometrics and Intelligent Laboratory Systems, 240, 104923. DOI: 10.1016/j.chemolab.2023.104923

[11] Tian, Z., Yu, X., Tang, F. (2025). Soft-sensing model of calcination zone temperature of rotary kiln based on principal component analysis and leaky integrator echo state network. Chemical Engineering Science, 309, 121472. DOI: 10.1016/j.ces.2025.121472

[12] Ren, L., Wang, T., Laili, Y., Zhang, L. (2021). A data-driven self-supervised LSTM-DeepFM model for industrial soft sensor. IEEE Transactions on Industrial Informatics, 18(9), 5859–5869. DOI: 10.1109/tii.2021.3131471

[13] Zhao, Y., Deng, X., Li, S. (2023). A nonlinear industrial soft sensor modeling method based on locality preserving stochastic configuration network with utilizing unlabeled samples. ISA Transactions, 139, 548–560. DOI: 10.1016/j.isatra.2023.04.012

[14] Gallareta, J.G., González-Menorca, C., Muñoz, P., Vidak Vasic, M. (2025). Advancements in soft sensor technologies for quality control in process manufacturing: A review. IEEE Sensors Journal, 25(9), 14575–14588. DOI: 10.1109/jsen.2025.3549596

[15] Sharman, M., Zweiri, Y., Jaradat, M., Al-Husari, R., Gan D., Lakmal, D. (2020). Deep-learning-based neural network training for state estimation enhancement: Application to attitude estimation. IEEE Transactions on Instrumentation and Measurement, 69(1), 24–34. DOI: 10.1109/tim.2019.2895495

[16] Nunes, F., Sousa, F. (2024). Deep Learning Soft-Decision GNSS Multipath Detection and Mitigation. Sensors, 24(14), 4663. DOI: 10.3390/s24144663

[17] Xu, O., Yang, Z., Ge, Z. (2025). Online quality estimation in chemical processes with random subspace deep partial least squares model. Chemical Engineering Science, 306, 1–10. DOI: 10.1016/j.ces.2025.121295

[18] Wang, Y., Shen, F., Ye, L. (2025). A knowledge-refined hybrid graph model for quality prediction of industrial processes. Engineering Applications of Artificial Intelligence, 139, 1–13. DOI: 10.1016/j.engappai.2024.109711

[19] Du, X., Yao, Y. (2024). FDA-SCN network based soft sensor for wastewater treatment process. Polish Journal of Environmental Studies, 33(1), 491–501. DOI: 10.15244/pjoes/169617

[20] Najafzadeh, M., Zahiri, A. (2015). Neuro-fuzzy GMDH-based evolutionary algorithms to predict low discharge in straight compound channels. Journal of Hydrologic Engineering, 20(12), 04015026. DOI: 10.1061/(asce)he.1943-5584.0001185

[21] Du, S, Gong, M., Wang, Q. (2023). Development of a NOx calculation model for low-speed marine diesel engines based on soft measurement technology. Applied Sciences, 13(11), 6373. DOI: 10.3390/app13116373

[22] Shahabi S., Khanjani, M., Kermani, M.H. (2016). Hybrid wavelet-GMDH model to forecast significant wave height. Water Supply, 16(2): 453–459. DOI: 10.2166/ws.2015.151

[23] Ebtehaj, I., Sammen, S., Sidek, L.M., Malik, A., Sihag, P., Mohammed, A., Chau, K., Bonakdari, H. (2021). Prediction of daily water level using new hybridized GS-GMDH and ANFIS-FCM models. Engineering Applications of Computational Fluid Mechanics, 15(1), 1343–1361. DOI: 10.1080/19942060.2021.1966837

[24] Harandizadeh, H., Armaghani, D., Khari, M. (2021). A new development of ANFIS–GMDH optimized by PSO to predict pile bearing capacity based on experimental datasets. Engineering with Computers, 35(4), 1213–1228. DOI: 10.1007/s00366-019-00849-3

[25] Sun, W.Z., Wang, J.S., Gao, S.Z. (2017). Soft-sensor modeling of PVC polymerizing process based on F-GMDH-type neural network algorithm. Journal of Sensors, 2017, Article ID 7872030. DOI: 10.1155/2017/7872030