Parameter Tuning and Modeling of a Rotary Kiln using Physics-Informed Neural Networks

Janak M. Patel [1] **Vishal Jadhav** [1] **Anirudh Deodhar** [1] **Shirish Karande** [1] **Venkataramana Runkana** [1]

Abstract

Physics-informed neural network models of industrial systems often fail to provide accurate predictions consistently over time because of temporal variations in raw material characteristics, plant operating and environmental conditions, and asset health. This necessitates updating of model parameters, which, in turn, requires exploration of the parameter space and identification of accurate values of model parameters regularly to enhance the reliability of models. To address this need, we present a sequential training and tuning methodology consisting of solving both forward and inverse problems of PINNs and parameter discovery via optimization. This methodology is tested for modeling of heat transfer in a rotary kiln, a common equipment in many process industries such as chemicals, steel, cement and materials. The proposed approach not only uncovers accurate model parameters but also helps in building a robust PINN model. Model predictions using parameters obtained through the proposed approach are in fairly good agreement with data from an industrial rotary kiln. This method can update model parameters as needed, offering more reliable and accurate predictions compared to traditional approaches.

1. Introduction

Rotary kiln, a cylindrical vessel rotating along its axis, is a common equipment in several process industries such as chemicals, pulp and paper, cement, minerals and metals, and food processing. It is essentially a multi-phase reactor in which reactions between gases and solids occur at high temperatures and result in the generation of products. One of the main challenges faced in the operation of these kilns is the formation of rings inside the kiln due to deposition of materials at certain locations. It happens due to various physico-chemical phenomena that take place inside the kiln. Due to this, the kiln production rate decreases and the quality of product deteriorates because the effectiveness of heat transfer between the materials diminishes(Runkana et al., 2010). In industrial systems like a rotary kiln, obtaining measurements is a complex task, and when available, they often tend to be sparse(Cai et al., 2021). However, health of the equipment changes with time due to material degradation and maintenance activities. Moreover, raw material characteristics, environmental conditions, and operating conditions of industrial processes also change significantly over time(Zagorowska et al., 2020). Because of these reasons, accuracy of physics-based models of rotary kilns that are commonly used for process simulation, analysis and optimization deteriorates and necessitates updating of appropriate model parameters such as heat transfer coefficients and reaction rate constants.

Physics-Informed Neural Networks (PINNs) have emerged as a promising alternative for solving physics-based models represented by nonlinear partial differential equations(Raissi et al., 2019). Recent developments emphasize the utility of PINNs for industrial systems for e.g. monitoring the health of diesel engines(Nath et al., 2023) and air preheaters(Jadhav et al., 2022). As mentioned above, parameters in model for industrial processes change with time. Since PINNs are incorporate physics-based models, such parameter changes can lead to deterioration of prediction accuracy of PINNs. Therefore, there is a need for a framework that can identify the parameters to be modified and update the parameters during training of the PINNs. Accordingly, the objective of this work is to develop a sequential training and tuning framework to tackle the aforementioned challenge for updating of models of industrial systems. We propose a sequential training and tuning methodology in which forward and inverse problems of the PINN are solved simultaneously, backed by parameter discovery through optimization. It was applied for modeling of heat transfer in a rotary kiln with an inert bed of particles, represented by a set of differential algebraic equations (DAEs). Additionally, the PINN was trained using conventional methods to demonstrate the efficacy of the proposed framework.

[1]TCS Research, Pune, Maharashtra, India. Correspondence to: Janak M. Patel <janak.mpatel@tcs.com>, Vishal Jadhav <vi.suja@tcs.com>.

Accepted at the ICML 2024 Workshop on AI for Science: Scaling in AI for Scientific Discovery

2. Methodology

2.1. Governing Differential and Algebraic Equations

A Pilot Rotary Kiln is being considered for this research work (Barr, 1986). Heat transfer between gases and solids is the main driving force for the reactions to occur in the kiln and it happens through multiple modes, namely, convection, conduction and radiation as shown in Figure 1. For simplification, the process is assumed to be at a steady-state. The energy balance equations for the gas and the solid phases are represented by set of ordinary differential equations (ODEs). Under steady-state conditions, the internal kiln wall temperature (T_w) and the external shell temperature (T_{sh}) are implicitly estimated through an energy balance around the kiln wall and the shell, represented by a set of algebraic equations (AEs). This leads to the formulation of a set of algebraic equations. Combining these two sets of equations results in a set of DAEs for modeling heat transfer in a rotary kiln, given by Equations (1)-(4) below.

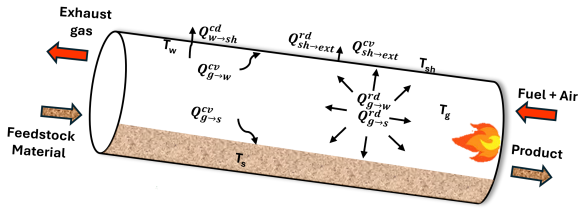

Figure 1. Schematic of a rotary kiln with various heat fluxes

$$\dot{m}_s C_{p,s} \frac{dT_s}{dx} = Q^{cv}_{g\to s} + Q^{cd}_{w\to s} + Q^{rd}_{g\to s} + Q^{rd}_{w\to s} \quad (1)$$

$$\dot{m}_g C_{p,g} \frac{dT_g}{dx} = Q^{cv}_{g\to s} + Q^{cv}_{g\to w} + Q^{rd}_{g\to s} + Q^{rd}_{g\to w} \quad (2)$$

$$Q_{w\to ext} = Q^{rd}_{g\to w} + Q^{cv}_{g\to w} - Q^{rd}_{w\to s} - Q^{cd}_{w\to s} \quad (3)$$

$$Q_{w\to sh} = Q^{rd}_{g\to w} + Q^{cv}_{g\to w} - Q^{rd}_{w\to s} - Q^{cd}_{w\to s} \quad (4)$$

Here, T_s and T_g are temperatures of solid and gas(K), $\dot{m}_s$ and $\dot{m}_g$ are the mass flow rates of solid and gas (kg/s), and $C_{p,s}$ and $C_{p,g}$ are specific heats capacities of the solid and gas phase$(J/kg.K)$,respectively. Q represents the heat flux per unit length of the kiln, with superscripts indicating the mode of heat transfer (convective(cv), radiative(rd), or conductive(cd)) and subscripts denoting the phases involved in the heat exchange. Here, $Q^{cv}_{g\to s}$ is rate of heat transfer by convection between gas and solid, $Q^{cd}_{w\to s}$ is rate of heat transfer by conduction between wall and solid, $Q^{rd}_{g\to s}$ is rate of heat transfer by radiation between gas and solid, $Q^{rd}_{w\to s}$ is rate of heat transfer by radiation between wall and solid, $Q^{cv}_{g\to w}$ is rate of heat transfer by convection between the gas and the wall, $Q_{w\to ext}$ is net heat transfer between wall and external environment, and $Q_{w\to sh}$ is net heat transfer between the wall and the shell. Additional information on calculating heat fluxes can be found in the appendix.

2.2. Physics-Informed Neural Network

2.2.1. Architecture of Physics-Informed Neural Network

The PINN framework for the rotary kiln is a fully connected deep neural network as illustrated in the Figure 2, where the input coordinate x yields the output $T(x)$, consisting of solid, gas, wall, and shell temperatures $([T_s, T_g, T_w, T_{sh}])$. The neural network comprises of multiple hidden layers, and propagation through each layer is governed by the following equation:

$$Y = \sigma(\boldsymbol{W}X + \boldsymbol{b}) \quad (5)$$

Here, X and Y are input and output of the neural network, respectively. $\boldsymbol{W}$ and $\boldsymbol{b}$ are trainable weights and bias, respectively. $\sigma(\cdot)$ is the activation function. Optimization is carried out to minimize both the losses associated with both the DAEs and the data. In this context, the data loss is computed using the dataset compiled by(Barr et al., 1989), available in the supplementary material by (Hanein et al., 2017) which includes sparse measurements of gas, solid, and wall temperatures for 9 experiments. The loss functions are expressed as

$$\mathcal{L}_{Total} = w_{DAE}\mathcal{L}_{DAE} + w_{data}\mathcal{L}_{data} \quad (6)$$

Here, $\mathcal{L}_{DAE}$ represents residual of the governing DAEs system, and $\mathcal{L}_{data}$ signifies the data loss calculated for sparse measurements. w_{DAE} denotes weight attributed to the DAEs loss, while w_{data} denotes weight assigned to data loss. The loss associated with boundary conditions is also incorporated into the data loss term.

Furthermore, we introduce two additional losses due to energy constraints , denoted as F^5_f and F^6_f, by leveraging ODEs and AEs. This is accomplished by subtracting Equation (1) from Equation (2) and, likewise, subtracting Equation (4) from Equation (3).The primary objective of incorporating these two energy constraints is to enhance convergence of the PINN. The expressions below are used to calculate the data and the DAEs losses:

$$\mathcal{L}_{data} = \frac{1}{N_s}\sum_{i=1}^{N_s}\left(T_{s_i} - \hat{T}_{s_i}\right)^2 + \frac{1}{N_g}\sum_{i=1}^{N_g}\left(T_{g_i} - \hat{T}_{g_i}\right)^2 + \frac{1}{N_w}\sum_{i=1}^{N_w}\left(T_{w_i} - \hat{T}_{w_i}\right)^2 \quad (7)$$

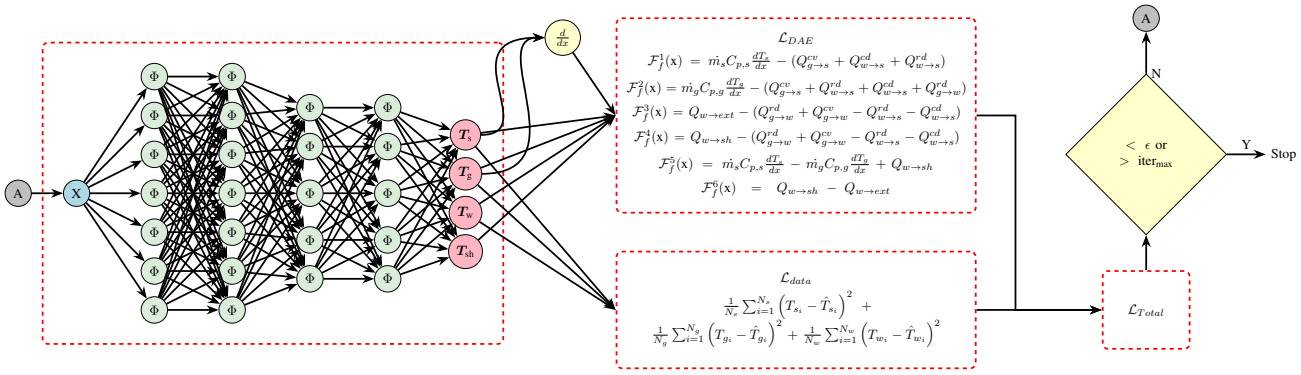

Figure 2. PINN Architecture

$$\mathcal{L}_{DAE} = \sum_{j=1}^{6} \frac{1}{N_d} \sum_{i=1}^{N_d} \left| \mathcal{F}_f^j(\mathrm{x_i}) \right|^2 \tag{8}$$

$$\mathcal{F}_f^1(\mathrm{x}) = \dot{m}_s C_{p,s} \frac{dT_s}{dx} - (Q_{g\to s}^{cv} + Q_{w\to s}^{cd} + Q_{g\to s}^{rd} + Q_{w\to s}^{rd}) \tag{9}$$

$$\mathcal{F}_f^2(\mathrm{x}) = \dot{m}_g C_{p,g} \frac{dT_g}{dx} - (Q_{g\to s}^{cv} + Q_{g\to w}^{cv} + Q_{g\to s}^{rd} + Q_{g\to w}^{rd}) \tag{10}$$

$$\mathcal{F}_f^3(\mathrm{x}) = Q_{w\to ext} - (Q_{g\to w}^{rd} + Q_{g\to w}^{cv} - Q_{w\to s}^{rd} - Q_{w\to s}^{cd}) \tag{11}$$

$$\mathcal{F}_f^4(\mathrm{x}) = Q_{w\to sh} - (Q_{g\to w}^{rd} + Q_{g\to w}^{cv} - Q_{w\to s}^{rd} - Q_{w\to s}^{cd}) \tag{12}$$

$$\mathcal{F}_f^5(\mathrm{x}) = \dot{m}_s C_{p,s} \frac{dT_s}{dx} - \dot{m}_g C_{p,g} \frac{dT_g}{dx} + Q_{w\to sh} \tag{13}$$

$$\mathcal{F}_f^6(\mathrm{x}) = Q_{w\to sh} - Q_{w\to ext} \tag{14}$$

The terms N_d, N_s, N_g, and N_w correspond to number of collocation points, count of locations for solid temperature, gas temperature, and wall temperature, respectively. Details of the neural network and related terms are provided in Table 1. Automatic differentiation is employed for evaluating the gradients of ODEs(Baydin et al., 2018).

2.2.2. Conventional Training

In the conventional training approach, PINN is trained by simultaneously minimizing all loss functions. This involves optimizing the network using both the data loss and the physics-based loss derived from the governing differential equations. The PINN for rotary kiln is trained using the conventional training approach for the purpose of comparison with the approach proposed in this work. Initially, the PINN is trained using the available sparse measurements. This preliminary training helps in initialization of the network weights. Following this, the PINN is further trained using both the data loss and physics (DAEs) loss, with untuned parameters of the DAEs system.

Table 1. Neural Network Configuration

Number of layers	4
Layer wise Neuron configuration	[64,64,32,32]
Activation Function	*tanh*
Optimizer	Adam
Collocation Points	1000
Loss Metric	Mean Sum Squared
System Discription	ubuntu:20.04, 20 GB RAM

2.2.3. Sequential Training and Tuning Strategy

Addressing the challenge of solving industrial systems requires exploring and discovering adjusted parameters to enhance model reliability. If sparse measurements of at least one temperature are available, we propose the following systematic step-by-step training and tuning strategy to tackle this challenge (Algorithm1):

Step 1 - Exclusive Training on Data Loss : In the first step, we utilize the available sparse measurements for at least one temperature. The PINN is exclusively trained using these sparse measurements, focusing solely on minimizing data loss with the DAEs loss weight set to zero. The primary objective of this step is to use the weights obtained from this model to initialize the next step, thereby accelerating the convergence process.

Step 2 - Training with Energy Constraints and Data

Loss : It is assumed that sparse measurements for at least one temperature are available. Solving all four energy constraints simultaneously using these measurements provides rough estimates for the remaining temperatures. Hence, in this step, the PINN is trained by minimizing both data loss and DAEs loss. However, Only the loss components of all four energy constraints ($\mathcal{F}_f^3$, $\mathcal{F}_f^4$, $\mathcal{F}_f^5$, $\mathcal{F}_f^6$) are activated as only energy constraints need to be solved. This training step generates rough predictions for all four temperatures. The accuracy of these predictions improves as sparse measurements for more temperatures become available.

Step 3 - Parameter Discovery via Optimization : In this step, the trained PINN model from step 2 predicts all four temperatures, and based on these predictions, we compute the DAEs loss for each equation. If the data loss is significantly lower than the DAEs loss, it indicates a close match between our predictions and the actual data, despite a high value of physics loss. In such cases, any remaining discrepancies in the DAEs system can likely be traced back to parameters that need tuning. Hence, we employ Particle Swarm Optimization (PSO) to fine-tune the DAE system parameters through an inverse problem approach, aiming to minimize the overall DAE loss. The optimization process considers a range of $\pm 5\%$ around the available parameter values as bounds. This step discovers the modified parameters of the physics model.

Step 4 - Comprehensive Training : In this final step, we utilize the modified parameters obtained from the previous optimization step. We train the model by simultaneously minimizing both data loss and DAE loss, activating all components of the DAE loss in this training phase.

Algorithm 1 Sequential Training and Tuning Strategy

Given: Measurements of one or more temperatures
$\mathcal{P} \leftarrow$ DAE Model Parameter
Initialize: $W_0, b_0 \leftarrow$ Parameter of the neural network
Step 1: Assign $w_{DAE} = 0, w_{data} = 1$
Train the neural network
Update $W_1, b_1 \leftarrow W_0, b_0$
Step 2: Assign $w_{DAE} = 1, w_{data} = 1$,
Activated DAE loss Component: F_f^3 to F_f^6
Train the neural network
Update $W_2, b_2 \leftarrow W_1, b_1$
Step 3: Predict $T \leftarrow feedforward(W_2, b_2)$
Calculate L_{DAE}
If $L_{DAE} >> L_{data}$ **then**
Fine Tune $\mathcal{P}$ such that L_{DAE} is minimum
Step 4: Assign $w_{DAE} = 1, w_{data} = 1$,
Activated DAE loss Component : All
Train the neural network
Update $W_3, b_3 \leftarrow W_2, b_2$

3. Result and Discussion

To demonstrate the effectiveness of the sequential training and tuning strategy, PINN was trained using both the conventional simultaneous training method and the sequential training and tuning approach.

- **Conventional Training:** Sparse measurements of gas, solid, and wall temperatures for nine different experiments are available for this pilot rotary kiln, which is well-documented in (Hanein et al., 2017). Additional details of the experiment are presented in Table 5 in the Appendix. For Experiment 3, the PINN was trained using the conventional simultaneous training method, in which both data loss and DAE loss are minimized simultaneously. The parameters of the physics model were taken from the literature (Hanein et al., 2017) and are listed in Table 2 under the untuned parameter column. To train the PINN for the rotary kiln, the model was initially trained for 30,000 epochs using only the sparse measurements available for gas, solid, and wall temperatures. Following this, the PINN undergoes further training for 150,000 epochs, utilizing all data and all components of DAEs loss. The training history of the PINN model using this conventional training method is shown in Figure 3. The model's predictions are compared with actual measurements in Figure 4.

  The loss value for each loss function is documented in Table 3. While the loss value for DAEs seems acceptable, the data loss is notably high, suggesting that the PINN may not have converged properly. The model provides moderate prediction results, as seen Figure 4. Since shell temperature measurements are not available, we can only validate them qualitatively. The shell temperatures should be the lower than all other temperatures and should increase from the feed end (0 m) to the burner end (5.5 m) as the kiln temperature increases in that direction. This trend is also visible in the plot, indicating that the shell temperature prediction is physically consistent.

- **Sequential Training and Tuning:** PINN for the rotary kiln was trained for experiment 3 using sequential training and tuning strategy. The outcomes of each step in the sequential training and tuning strategy are presented here. It's important to note that all plots shown here correspond specifically to experiment number 3. In the first step, the model undergoes exclusive training using data loss for 30,000 epochs, as detailed in Section 2. The predictions from this model are depicted in Figure 5. The model accurately predicts solid and gas temperatures. However, due to the lack of available data for shell temperature, its predictions exhibit random behavior in this regard, lacking theoretical

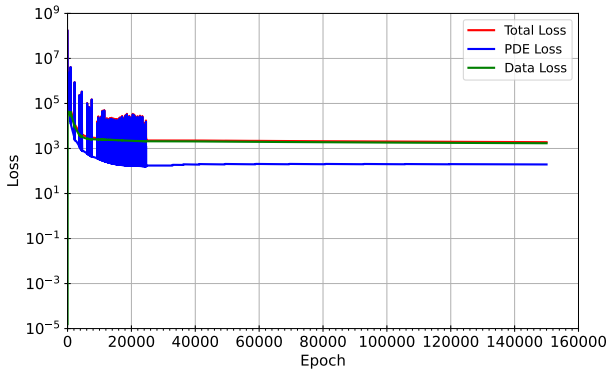

Figure 3. Conventional Training : Loss values w.r.t Epochs

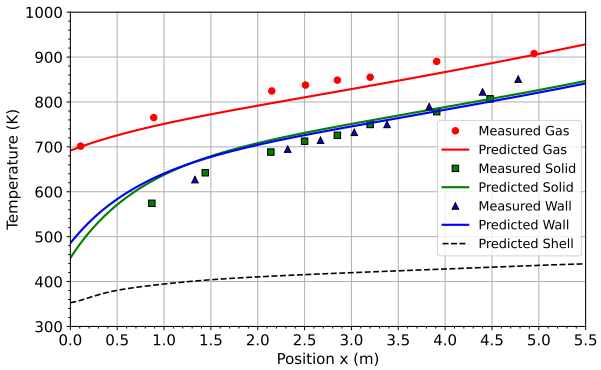

Figure 4. Conventional Training: Comparison of Measured and Predicted Temperatures

coherence. Furthermore, the prediction of wall temperature at the feedend is inaccurate due to the absence of measurements.

In the second step, the model undergoes 30,000 epochs of training, incorporating both energy constraints and data loss. The resulting predictions are shown in Figure 6. The model predicts all four temperatures accurately. Shell temperature increases towards the burner end, as expected due to the higher temperature at that end compared to the feed end. Given its continuous contact with ambient conditions, shell temperature is anticipated to be the lowest. This trend aligns with expectations and provides qualitative validation since measurements of shell temperature are unavailable. The individual loss values for each component are detailed in Table 3. It's evident that the reported DAEs loss is significantly higher than the data loss, yet it matches well with experimental data. This clearly indicates the necessity for parameter tuning. We identified the parameters that needed fine-tuning and discovered the modified parameters, which are detailed in Table 2.

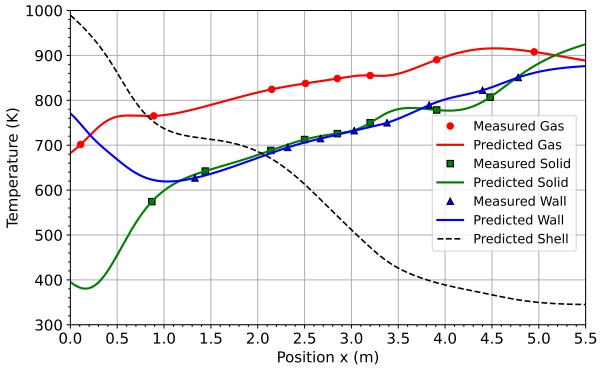

Figure 5. Sequential Training - Step 1 Exclusive Trained on Data Loss Model: Comparison of Measured and Predicted Temperatures

Next, the PINN was trained for 150,000 epochs using both losses. The training history for step 4 is depicted in Figure 7. Comparisons between gas, solid, and wall temperature predictions from the PINN model and the experimental data are shown in Figure 8, indicating reasonably accurate predictions. Additionally, the trend in shell temperature prediction aligns with expectations. Therefore, the PINN model provides reasonably accurate predictions for all four temperatures. Detailed breakdown of the loss of each component of the DAE system and data at each step is shown in Table 3. The loss values for both DAEs and data are notably low after step 4, indicating that the PINN has converged effectively.

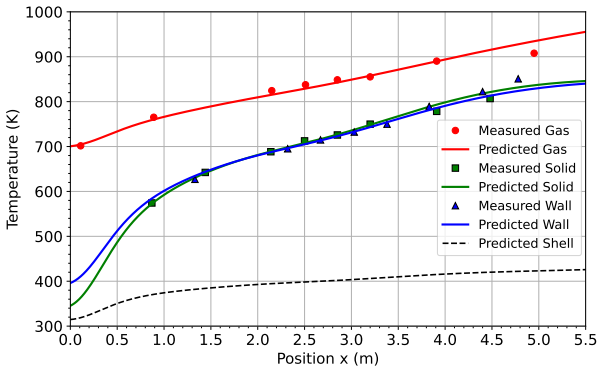

Figure 6. Sequential Training - Step 2 Trained with Energy Constraint and Data Loss Model: Comparison of Measured and Predicted Temperatures

The comparison of loss values after conventional training and sequential training and tuning is presented in Table 2.

Table 2. Untuned and Fine-tuned value of Parameters (in bold faced)

Parameter Name	**Untuned Value**	**Fine Tuned Value**
Thermal Conductivity of Air	$\mathbf{0.0002} * T^{\mathbf{0.8218}}$	$\mathbf{0.000138} * T^{\mathbf{0.8218}}$
Emissivity of Air	$\mathbf{10^{1.9}}/T_g$	$\mathbf{10^{1.97}}/T_g$
Emissivity of Solid	**0.9**	**0.81**
Emissivity of Wall	**0.85**	**0.935**
Emissivity of Shell	**0.8**	**0.95**
Air volumetric expansion coefficient	$\mathbf{987.24} * T^{\mathbf{-0.996}}$	$\mathbf{978.16} * T^{\mathbf{-0.995}}$
Density of Air	$\mathbf{352.94} * T^{\mathbf{-0.1}}$	$\mathbf{359.16} * T^{\mathbf{-0.94}}$
Dynamic Viscosity of Air	$(\mathbf{-10} * \mathbf{10^{-12}})T^2 + (\mathbf{5} * \mathbf{10^{-8}})T + (\mathbf{4} * \mathbf{10^{-6}})$	$(\mathbf{-8.6} * \mathbf{10^{-12}})T^2 + (\mathbf{5.99} * \mathbf{10^{-8}})T + (\mathbf{3.488} * \mathbf{10^{-6}})$
Quartz Heat Capacity	Different set of coefficients for different temperature ranges(Haas Jr et al., 1981)	$(\mathbf{-7.77338} * \mathbf{10^{2}}/T_s) + \mathbf{83.2101} + (\mathbf{1.09962} * \mathbf{10^{-2}}T_s^2/2)$

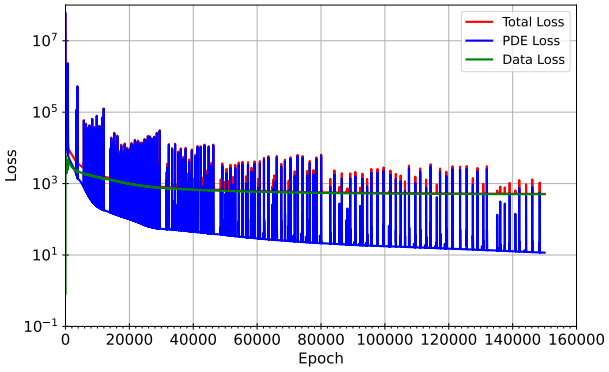

Figure 7. Sequential Training - Step 4 Comprehensive Training : Loss values w.r.t Epochs

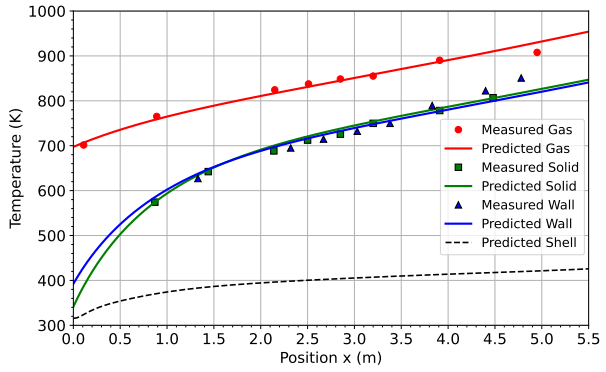

Figure 8. Sequential Training - Step 4 Comprehensive Trained Model : Comparison of Measured and Predicted Temperatures

Table 3. MSE loss values of the each component after training

Loss component		**Conventional Training**	**Sequential Training and Tuning**		
			Step 1	**Step 2**	**Step 4**
DAE loss	$\mathcal{F}_f^1$	98.33	2.31×10^6	7.85×10^5	1.24
	$\mathcal{F}_f^2$	15.08	6.24×10^6	7.85×10^5	0.47
	$\mathcal{F}_f^3$	13.32	8.63×10^6	10.68	2.67
	$\mathcal{F}_f^4$	13.49	1.09×10^7	10.46	1.25
	$\mathcal{F}_f^5$	7.01	1.44×10^7	2.71	0.78
	$\mathcal{F}_f^6$	47.89	1.54×10^7	9.29	5.21
	Total	195.12	5.80×10^7	1.57×10^6	11.63
Data loss		1709.76	0.8262	351.53	503.57

Table 4. Comparison of Sequential PINN model with Conventional PINN model and Numerical Simulation across all nine experiments

	Maximum error (K)			**Mean error (K)**		
	T_s	T_g	T_w	T_s	T_g	T_w
Sequential PINN	44.16	47.84	39.38	17.79	18.59	13.74
Conventional PINN	279.79	213.09	249.86	172.15	132.09	155.44
Numerical Simulation (Hanein et al., 2017)	37.8	54.1	39.6	13.9	15.5	13.5

The DAE loss from conventional training is almost ten times higher than that from sequential training and tuning. Additionally, the data loss is significantly higher for conventional training compared to sequential training and tuning. This observation distinctly reveals that the sequential training and tuning strategy provides enhanced convergence compared to the conventional training method. Our method fine-tunes the parameters and effectively enhances the convergence of the PINN for the rotary kiln.

In a similar vein, dedicated PINN models for the remaining eight experiments are developed using both training approaches. A summary of these PINN models is presented in Table 4, which includes statistics on the deviation of the PINN's temperature predictions from the experimental results across all nine experiments, while comparison of the Sequential PINN model and the Conventional PINN model for individual experiments is presented in Table 6 in the Appendix. Additionally, these results are compared with numerical simulations documented in the literature(Hanein et al., 2017). The maximum error denotes the largest difference between model predictions and experimental measurements, while the mean error represents the average difference between predictions and experimental measurements across all nine experiments. Upon comparison, it becomes evident that PINN trained with conventional training exhibit approximately 10 times higher mean error than the PINN trained with the sequential training and tuning method. Additionally, the maximum error is 5 times greater in conventional training than in sequential training. This clearly indicates that our method outperforms the conventional training method. Therefore, our proposed sequential training and tuning method results in more accurate and reliable PINN models than conventional training. Furthermore, the accuracy of these PINN models is closely aligned with the simulation results from the rigorous physics-based model. It is well-known that the inference time of PINN is significantly shorter than that of numerical solutions; hence, PINN developed using our proposed method can be employed for real-time predictions.

4. Conclusions

Models for industrial processes have to be updated regularly to account for the time varying behavior of quality of raw materials, health of equipment, environment and plant operating conditions. A sequential training and tuning methodology was proposed to address the need for discovering and fine tuning of model parameters in a PINN. This methodology was applied for modeling of heat transfer in a rotary kiln, commonly employed in diverse process manufacturing industrie. This methodology identifies modified parameters of the system and develops a PINN with finely tuned parameters. The comparative analysis strongly emphasizes that our proposed method offers more reliable, accurate predictions, and improves convergence compared to the conventional training method. Moreover, predictions of PINN obtained through the proposed methodology are in reasonable agreement with data from a pilot scale rotary kiln. Since the inference time for a PINN model is quite low compared to simulation using a rigorous physics-based model, the proposed approach is expected to be suitable for real-time applications in industrial environments.

Acknowledgements

The authors thank the management of Tata Consultancy Services Limited to publish this article and Dr. Harrick Vin, Mr. K. Ananth Krishnan and Dr. Gautam Shroff for their encouragement and support.

References

Barr, P., Brimacombe, J., and Watkinson, A. A heat-transfer model for the rotary kiln: Part i. pilot kiln trials. *Metallurgical and Materials Transactions B*, 20:391–402, 1989.

Barr, P. V. *Heat transfer processes in rotary kilns*. PhD thesis, University of British Columbia, 1986.

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18(153): 1–43, 2018.

Boateng, A. A. *Rotary kilns: transport phenomena and transport processes*. Butterworth-Heinemann, 2015.

Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6):060801, 2021.

Haas Jr, J. L., Robinson Jr, G. R., and Hemingway, B. S. Thermodynamic tabulations for selected phases in the system cao-al2o3-sio2-h2 at 101.325 kpa (1 atm) between 273.15 and 1800 k. *Journal of Physical and Chemical Reference Data*, 10(3):575–670, 1981.

Hanein, T., Glasser, F. P., and Bannerman, M. N. One-dimensional steady-state thermal model for rotary kilns used in the manufacture of cement, Apr 2017. URL https://tandf.figshare.com/articles/dataset/One-dimensional_steady-state_thermal_model_for_rotary_kilns_used_in_the_manufacture_of_cement/4810927/2.

Hottel, H. and Sarofim, A. *Radiative Transfer*. McGraw-Hill, New York, 1967.

Jadhav, V., Deodhar, A., Gupta, A., and Runkana, V. Physics informed neural network for health monitoring of an air preheater. In *PHM Society European Conference*, volume 7, pp. 219–230, 2022.

Lehmberg, J., Hehl, M., and Schügerl, K. Transverse mixing and heat transfer in horizontal rotary drum reactors. *Powder Technology*, 18(2):149–163, 1977.

Liu, Z., Meng, P., Liang, Y., Li, J., Miao, S., and Pan, Y. Research on lime rotary kiln temperature prediction by multi-model fusion neural network based on dynamic time delay analysis. *Thermal Science*, (00):264–264, 2023.

Moya, C. and Lin, G. Dae-pinn: a physics-informed neural network model for simulating differential algebraic equations with application to power networks. *Neural Computing and Applications*, 35(5):3789–3804, 2023.

Mujumdar, K. and Ranade, V. Simulation of rotary cement kilns using a one-dimensional model. *Chemical engineering research and design*, 84(3):165–177, 2006.

Nath, K., Meng, X., Smith, D. J., and Karniadakis, G. E. Physics-informed neural networks for predicting gas flow dynamics and unknown parameters in diesel engines. *arXiv preprint arXiv:2304.13799*, 2023.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.

Runkana, V., Natekar, P., Bandla, V., and Pothal, G. Mathematical modelling of sponge iron production in a rotary kiln. *SteelTech*, 5:9–17, 2010.

Stiasny, J., Chatzivasileiadis, S., and Zhang, B. Solving differential-algebraic equations in power systems dynamics with neural networks and spatial decomposition. *arXiv preprint arXiv:2303.10256*, 2023.

Tscheng, S. H. *Convective heat transfer in a rotary kiln*. PhD thesis, University of British Columbia, 1978.

Zagorowska, M., Wu, O., Ottewill, J. R., Reble, M., and Thornhill, N. F. A survey of models of degradation for control applications. *Annual Reviews in Control*, 50:150–173, 2020.

A. Description of Rotary Kiln

The rotary kiln is a cylindrical vessel that rotates around its axial position. Solid material is introduced at the upper end, while hot gas is fed from the lower end. The solid material undergoes heating as the gas moves from the lower to the upper end. In this pilot rotary kiln, granular material is introduced from one end, while natural gas firing is executed from the opposite end through burner arrangements as shown in Figure 1. The rotary kiln has a length of 5.5 meters and an internal diameter of 401 mm. The refractory lining, with a thickness of 93 mm, has a thermal conductivity of 0.4 W/m.K. The kiln shell was constructed using steel with a thickness of 6 mm, and its thermal conductivity is taken to be 45.2 W/m.K. The following assumptions are mode for developing the model(Hanein et al., 2017):

1. The bed of particles within the kiln is assumed to be unreactive and consists of silica
2. Chemical compositions of solids and gas phases are considered to remain constant along the length of the kiln, as there is no mass change between phases
3. Pressure drop is neglected along the length of the kiln, resulting in a constant gas velocity
4. A relatively constant bed height is assumed along the length of the kiln
5. The effect of axial temperature gradient driving axial conduction is ignored in the analysis

A.1. Heat Flux Computation: Analytical Evaluation

The simplified calculation for heat fluxes, crucial for evaluating ordinary differential and algebraic equations, is presented here. This particular computation is explained in more detail in the work by (Hanein et al., 2017).

Conduction ($\mathbf{Q^{cd}_{w\to s}}$)
Heat transfer through conduction mode occurs only between the solid and the wall. This heat flux between the solid bed and the wall can be calculated using the following expression.

$$Q^{cd}_{w\to s} = h^{cd}_{cw-s} P_{cw-s}(T_w - T_s)$$

where h^{cd}_{cw-s} is heat transfer coefficient and P_{cw-s} is perimeter of wall in contact with solid. The detailed explanation of calculating the heat transfer coefficient can be found((Lehmberg et al., 1977; Barr et al., 1989)).

Convection ($\mathbf{Q^{cv}_{g\to s}}, \mathbf{Q^{cv}_{g\to w}}$)
In a rotary kiln, convective heat transfer takes place between the gas and the solid material, as well as between the gas and the kiln wall. The corresponding heat fluxes for these convection phenomena can be computed using the following expressions

$$Q^{cv}_{g\to s} = h_{g-s} P_{g-s}(T_g - T_s)$$
$$Q^{cv}_{g\to w} = h_{g-w} P_{g-w}(T_g - T_w)$$

Here, h_{g-s} represents convective heat transfer coefficient between gas and solid, and h_{g-w} represents the convective heat transfer coefficient between gas and wall (Tscheng, 1978).P_{g-s} and P_{g-w} denote perimeters of the exposed bed and exposed wall, respectively.

Radiation ($\mathbf{Q^{rd}_{g\to s}}, \mathbf{Q^{rd}_{g\to w}}, \mathbf{Q^{rd}_{w\to s}}$)
Radiative heat transfer takes place among gas to solid, gas to wall, and wall to solid within the rotary kiln. The corresponding expressions are as follows(Hottel & Sarofim, 1967):

$$Q^{rd}_{g\to s} = \sigma(\varepsilon_s + 1) P_{g-s} \frac{\varepsilon_g T_g^4 - \alpha_s T_s^4}{2}$$

$$Q^{rd}_{g\to w} = \sigma(\varepsilon_w + 1) P_{g-w} \frac{\varepsilon_g T_g^4 - \alpha_w T_w^4}{2}$$

$$Q^{rd}_{w\to s} = \frac{\sigma(T_w^4 - T_s^4)}{(1-\varepsilon_w)/\varepsilon_w P_{s-w} + 1/F_{s-w} P_{w-s} + (1-\varepsilon_s)/\varepsilon_s P_{s-w}}$$

Table 5. Operating conditions of the Barr pilot kiln experiments(Barr et al., 1989)

Experiment number	Fuel flow rate (l/s)	Air flow rate(kg/s)	Solid mass flow rate(kg/s)	Particle diameter (m)
1	0.83	0.03454	0.01722	0.0025
2	1.02	0.06995	0.01722	0.0025
3	1.42	0.07105	0.01722	0.0025
4	1.97	0.07399	0.01722	0.0025
5	0.68	0.03577	0.01611	0.00058
6	0.9	0.0532875	0.01722	0.00058
7	1.04	0.075215	0.0175	0.00058
8	2	0.07215	0.01778	0.00058
9	2.53	0.075705	0.01805	0.00058

Here, σ is the Stefan-Boltzmann constant, whereas ε_s and ε_w stand for the emissivities of the solid and the wall, respectively. Similarly, α_s and α_w represent the absorptivities of the solid and the wall, respectively. $Fs - w$ represents the bed to wall form factor, and $Pw - s$ is perimeter of the exposed wall.

Heat loss from the kiln ($\mathbf{Q_{w\rightarrow ext}}$, $\mathbf{Q_{w\rightarrow sh}}$)
The heat loss from the kiln can be computed using the following equations(Hanein et al., 2017):

$$Q_{w\rightarrow ext} = \frac{T_w - T_{ext}}{R_{Total}}$$

$$Q_{w\rightarrow sh} = \frac{T_w - T_{sh}}{\sum_j R^{cd}_{wall,j}}$$

where, R_{Total} represents the cumulative resistance resulting from conduction across the kiln layers ($R^{cd}wall, j$), external convective resistance from the outer shell to the surroundings ($R^{cv}sh \rightarrow ext$), and external radiative resistance from the outer shell to the surroundings ($R^{rd}_{sh\rightarrow ext}$), as follows:

$$R_{Total} = \sum_j R^{cd}_{wall,j} + R^{cv}_{sh\rightarrow ext} + R^{rd}_{sh\rightarrow ext}$$

More details for calculating individual resistances can be found in (Hanein et al., 2017).

Table 6. Experiments wise comparison of Sequential PINN model with Conventional PINN model

Experiment Number		Maximum error(K)			Mean error(K)		
		T_s	T_g	T_w	T_s	T_g	T_w
1	**Sequential PINN**	20.04	37.13	24.47	12.02	19.11	10.14
	Conventional PINN	158.02	201.11	145.01	142.68	110.04	129.29
2	**Sequential PINN**	22.29	19.57	34.09	14.07	9.61	13.36
	Conventional PINN	145.02	130.18	137.66	141.87	110.25	126.46
3	**Sequential PINN**	12.43	22.15	39.38	5.72	7.27	15.93
	Conventional PINN	168.27	189.33	166.79	164.80	147.76	153.98
4	**Sequential PINN**	22.84	31.64	24.64	14.78	15.80	15.74
	Conventional PINN	213.21	211.28	212.11	201.93	168.08	201.58
5	**Sequential PINN**	44.16	36.42	22.55	25.66	27.33	7.53
	Conventional PINN	163.86	182.64	126.14	150.39	100.25	117.19
6	**Sequential PINN**	41.90	47.84	18.27	22.38	27.15	9.09
	Conventional PINN	162.43	198.89	135.51	152.15	115.55	125.63
7	**Sequential PINN**	40.61	37.34	26.19	22.22	17.72	11.92
	Conventional PINN	159.84	125.23	140.47	143.39	98.31	123.14
8	**Sequential PINN**	80.84	46.25	41.72	30.95	28.49	26.04
	Conventional PINN	221.00	178.18	214.59	210.68	150.63	201.65
9	**Sequential PINN**	69.21	36.54	46.96	25.70	26.26	22.49
	Conventional PINN	279.78	213.01	249.86	241.38	187.94	220.06