



## Modelling a Cement Precalciner by Machine Learning Methods

---

Amila Chandra Kahawalage and M.H. Wathsala N. Jinadasa

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

December 20, 2021

# Modelling a Cement Preheater by Machine Learning Methods

Amila C. Kahawalage, M.H. Wathsala N. Jinadasa

Department of Process, Energy and Environmental Technology, Faculty of Technology,  
Natural Science and Maritime Sciences, University of South-Eastern Norway, Kjølnes Ring 56, 3918, Porsgrunn, Norway  
{amila.c.kahawalage,Wathsala.jinadasa}@usn.no

## Abstract

This work is a feasibility study of modelling the calcination process in a cement preheater by employing machine learning algorithms. Calcination plays a significant role in characterising the clinker quality, energy demand and CO<sub>2</sub> emissions in a cement production facility. Due to the complex nature of the calcination process, it has always been a challenge to reasonably model the preheater system. This study is an attempt of finding a feasible alternative to answering this challenge. In this study, six machine learning algorithms were tested to analyse three output variables, which are, 1). the apparent degree of calcination, 2). CO<sub>2</sub> molar fraction (dry basis) and 3). water molar fraction in the preheater outlet stream. Fifteen input variables were used to train the algorithms, of which the values were obtained through a large number of simulated datasets by applying mass and energy balance to the preheater system. A number of machine learning algorithms showed better predictability and Artificial neural network (ANN) showed the best performance for all three output variables.

*Keywords: preheater, cement manufacturing process, machine learning, degree of calcination*

## 1 Introduction

### 1.1 Cement manufacturing and calcination

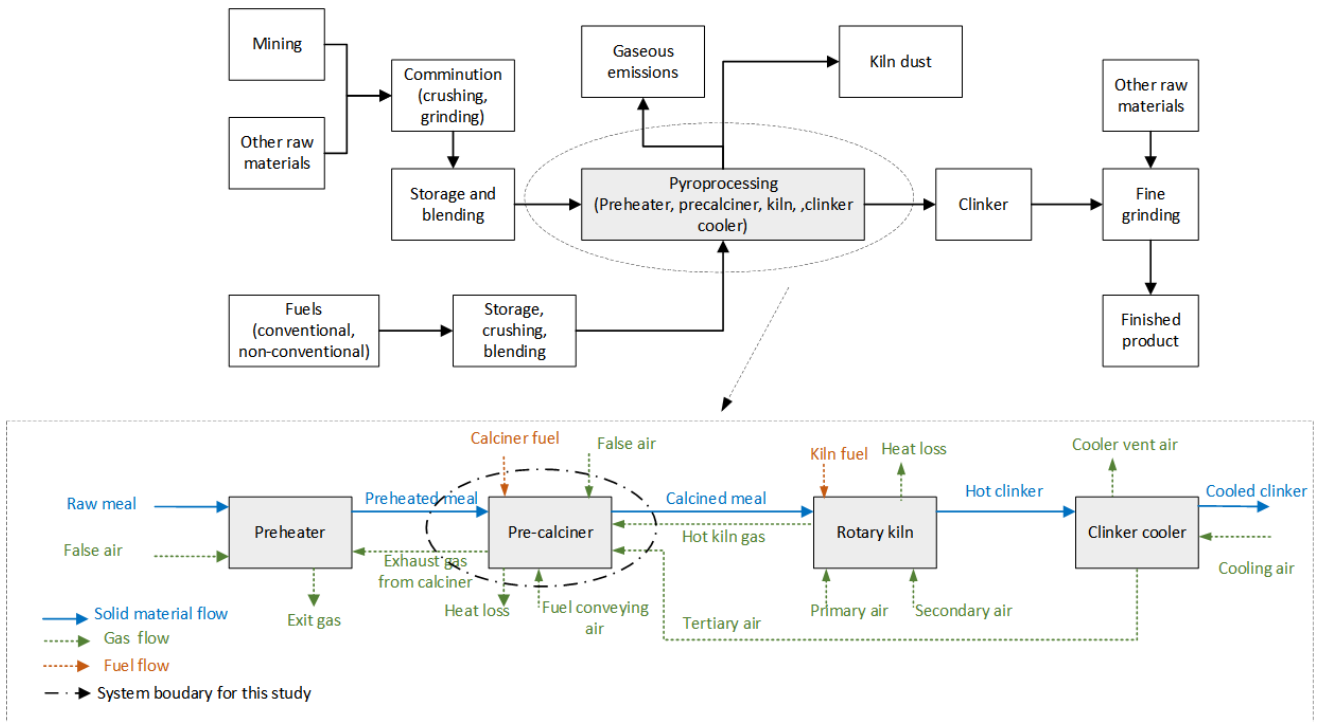
Cement is one of the frequently utilized materials in building infrastructure facilities. Cement manufacturing is a globally crucial industrial sector that is highly energy intensive. It is responsible for a considerable share of global CO<sub>2</sub> emissions. The dominant uses of carbon-intensive fuels, such as coal in clinker making and calcination process, are accountable for a large amount of CO<sub>2</sub> emissions in the cement industry. Calcination in the cement manufacturing process is a complex industrial phenomenon involving mass transfer, heat transfer, and physical and chemical reactions. Materials are subjected to high temperatures so as to cause a chemical and physical change. Process emissions from calcination of limestone are 60%, where 0.5 tonnes of CO<sub>2</sub> is emitted per tonne of clinker

production (IEA, 2008). The endotherm reaction at 950 °C in the calciner demands about 1700 MJ/t clinker energy, which is around 50% of total energy (WWFI, 2008).

Figure 1 shows a schematic diagram of a typical dry-based cement manufacturing facility. Most of the modern cement facilities are equipped with a preheater system located between the preheater and the rotary kiln. In the production process, raw materials, typically 80-90% limestones, are prepared by crushing, grinding and adding chemicals. This preprocessed raw material (which is referred a 'raw meal') is then preheated to 750°C and sent to the preheater (also called as calciner). Preheater initiates the chemical decomposition of limestone (CaCO<sub>3</sub>) into lime (CaO) and carbon dioxide (CaCO<sub>3</sub> ↔ CaO + CO<sub>2</sub>). About 90% of raw meal is calcined at this unit (GmbH, 2016). Preheater system provides direct combustion through solid-gas heat exchange, where it disperses and suspends cement raw meal powder in an airflow. The pre-calcined meal then enters the rotary kiln, where the remaining calcination process is completed. Clinker formation takes place in the kiln and finally the clinker is sent to the clinker cooler.

Stability and the effectiveness of the calcination process directly affects the final clinker quality, smooth operation in the subsequent rotary kiln operation and the energy requirement of the pyroprocessing unit. The exothermic process of fuel combustion and the endothermic process of carbonate decomposition in the raw meal occurs simultaneously in the preheater. The optimum operation of preheater conserves energy and reduces emissions associated with both preheater and rotary kiln. Calcination degree, which is an indicator of the performance of the preheater, is affected by several parameters such as temperature inside the calciner, residence time of the raw meal in the system, solid gas separation, dust circulation effect and kinetic behavior of raw materials (Mikulčić et al., 2012; Mohammadhadi, 2018).

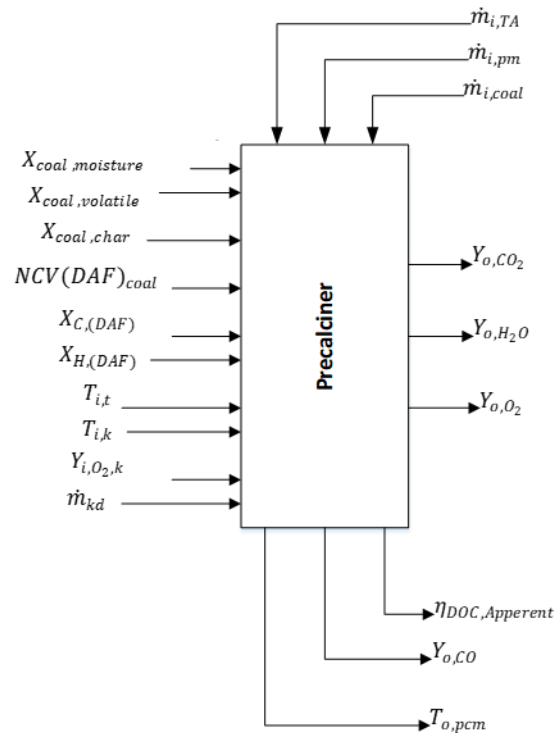
Calcination degree is expressed in two ways; either *true calcination degree* or *apparent calcination degree* (Tokheim, 1999). The apparent degree of calcination  $\eta_{app}$  is mentioned as ADOC in this paper, which is used as an indicator to monitor the calcination process in the



**Figure 1.** Schematic diagram of a typical cement manufacturing facility with details about input/output flow streams for the pyroprocessing section

cement production line because accurate calcination degree cannot be measured easily. However, it is not easy to measure the apparent degree of calcination online. Instead, samples are extracted from the process line and analyzed offline in the laboratory. The frequency between two subsequent analyses can be one hour or even several hours, depending on the availability of laboratory capacity. The precalciner outlet temperature is therefore used as the primary controlled variable in the precalciner to control the degree of calcination. Oxygen and carbon monoxide levels are also controlled because they are indicators of the fuel combustion and stabilisation of the process, respectively (Osmic et al., 2020)

Figure 2 shows different input and output variables to the cement precalciner system. These variables belong either to basic input streams (i.e. preheated raw meal, fuel and tertiary air) or primary output stream (i.e. calcined meal). Description of symbols can be found in Table 1. Some of these variables can be measured online by appropriate sensors, while others are difficult to measure. In such situations, they are computed using available measurements. The computation can be accomplished on the basis of appropriate assumptions



**Figure 2.** Input, performance and intermediate variables for a precalciner system

## 1.2 A modelling approach to assessing the performance of the precalciner

Several researchers have attempted to develop relationships between variables in precalciner process using the soft and hard modelling approach. Coupling, time-varying delay, and nonlinearity of precalciner system make it hard to establish an exact mathematical model to realize performance indicators such as ADOC. Mass and energy balance (MEB) provides a fundamental approach to derive correlation to determine a required process output. Authors in this study have experience in employing MEB to model precalciner. When there are input parameters which are unknown or cannot be measured directly, an iterative procedure is used during the MEB calculation. An example of an alternative approach to MEB is machine learning methods where this iterative process can be skipped.

Machine learning (ML) has shown promising results in modelling complex and nonlinear manufacturing processes that deal with noisy, limited and non-integrated data. Machine learning algorithms such as Support Vector Machine (SVM) and Artificial Neural Network (ANN) have proven their capabilities in this regard. (Gang and Hui, 2010) developed a model by using Least Squares Support Vector Machine (LS-SVM) with radial basis function (RBF) kernel for determining the apparent degree of calcination. The furnace temperature and pressure, the outlet temperature and pressure of the calciner, the temperature of the tertiary air and the lay-off quantity of cement raw were used as inputs to the model. (Griparis et al., 2000) proposed, adaptive, robust and fuzzy control to achieve the desired degree of precalcination of the raw meal, low carbon monoxide, while stabilising the precalcination process considering the multivariable dependencies in the precalciner system. (Yang et al., 2010) developed a back-propagation neural network (BPNN) and Radial basis function neural network (RBFNN) to assess the kiln temperature and oxygen content based on five variables which are coal flow to the kiln, coal flow to the precalciner, raw meal flow, rotary speed of kiln and negative pressure of the preheater exit.

The performance of the machine learning algorithms depends highly on the quality of input data. Therefore, collection and preparation of training dataset is an important step in the modelling process. Training data can be provided in three ways; 1) simulated data 2) actual process data and 3) designed experiment data. Simulated data is generated by theoretical models such as statistical models and computer simulations. Actual process data are randomly selected raw process data and many manufacturing companies have historical data in their database. Designed experimental data can be obtained using a Taguchi or Design of Experiment (DOE) approach. Among these three approaches, training and optimizing a model using a large number of less expensive simulation data and testing the model

with a smaller dataset of process data is a cost-effective approach.

This feasibility study aims to provide an alternative approach to conventional mass and energy balance to model precalciner in a cement manufacturing process. Simulated data from MEB calculations were used to train, validate, and test different machine learning models to predict apparent calcination degree, molar fraction of water and CO<sub>2</sub> (dry basis) in precalciner output. They were assessed based on known values of fifteen input variables.

## 2 Materials and methods

### 2.1 Input and output data

The first phase of modelling work in this study was selecting input and output variables for models. These variables are listed in Table 2, including their maximum, minimum, mean and standard deviations. The dataset included 20543 samples. The full-factorial design approach, a famous experiment design, was used to generate the synthetic input data matrix. These data were used to obtain the output data matrix by applying mass and energy balance to the precalciner. The system boundary of the model is shown in Figure 1.

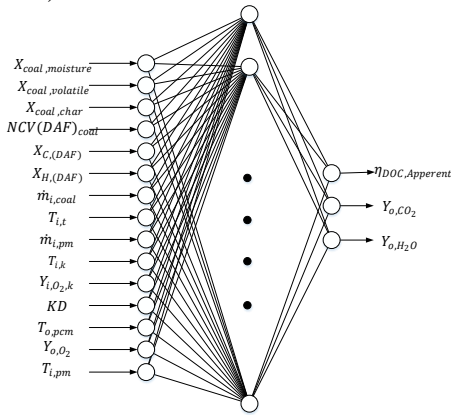
### 2.2 Methodology

**Table 1.** List of regression algorithms used to train models

Regression Model category	Regression model type
Linear regression model	1). Classical linear 2). Interaction linear 3). Robust linear 4). Stepwise linear
Regression trees	1). Fine Tree 2). Medium tree 3). Coarse Tree
Support vector machines (SVM)	1). Linear SVM 2). Quadratic SVM 3). Cubic SVM 4). Fine Gaussian SVM 5). Medium Gaussian SVM 6). Coarse Gaussian SVM
Gaussian Process Regression	1). Rational quadratic 2). Squared exponential 3). Matern 4). Exponential
Ensemble of Regression Trees	1). Boosted tree 2). Bagged tree
Artificial neural network	1). Levenberg-Marquardt back-propagation

All data were normalized before feeding to the models. Regression Learner App available in Matlab 2019

software was used for model development (Mathworks, 2021).



**Figure 3.** Model network architecture for ANN; The same inputs and outputs were also used for other regression models

Table 1 shows the list of algorithms that were used to train the dataset. These algorithms fall under six regression model categories as linear regression, SVM, regression tree, ensemble gaussian process regression (GPR) and ANN. There are different algorithm types (19 in total) under each of first five categories. The dataset was trained for all these algorithms. For ANN, the data was trained with Levenberg-Marquardt back-propagation algorithm with 20 neurons and one hidden

layer. Selection of number of neurons and hidden layers for the ANN model was an arbitrary option. Figure 3 illustrates the ANN network architecture representing inputs and outputs.

Three statistical indicators were used for evaluation of the model performance, which include mean absolute error (MAE), root mean squares error (RMSE) and coefficient of determination ( $R^2$ ). They were calculated as shown in Equation 1 to 3.

$$MAE = \sum_{i=1}^m \frac{(y_i - \hat{y}_i)}{m} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \left( \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \right) \quad (3)$$

Here  $\hat{y}_i$  is the estimated value by the model,  $y_i$  is the actual value of the response process (MEB based simulation data), and  $m$  is the number of samples in the dataset.

**Table 2.** Description of input and output variables to the precalciner model

No:	Input/output name	Symbol	Variable type	Unit	Minimum	Maximum	Mean	Standard deviation
1	Moisture content of the fuel	$X_{coal, moisture}$	Input	%	0.87	2.27	1.38	0.61
2	Volatile content of the fuel	$X_{coal, volatile}$	Input	%	16.17	31.79	24.47	7.15
3	Char content of the fuel	$X_{coal, char}$	Input	%	18.84	47.43	34.52	13.40
4	*NCV of coal	$NCV(DAF)^{coal}$	Input	J/kg	31.47	34.63	33.27	1.34
5	The carbon content of coal (DAF basis)	$X_C(DAF)$	Input	%	86.48	93.47	89.98	3.05
6	Hydrogen content of coal (DAF basis)	$X_H(DAF)$	Input	%	3.93	5.19	4.62	0.59
7	Coal mass flow rate	$\dot{m}_{i,coal}$	Input	kg/s	2.67	3.61	3.07	0.38
8	Tertiary air temperature	$T_{i,t}$	Input	K	742.70	908	821.67	67.32
9	Pre-heated meal mass flow rate	$\dot{m}_{i,pm}$	Input	kg/s	52.37	66	60.14	5.49
10	Pre-heated meal temperature	$T_{i,pm}$	Input	K	950	989	968.92	19.54
11	Kiln gas temperature	$T_{i,k}$	Input	K	1215	1250	1232.10	17.50
12	Kiln gas O <sub>2</sub> molar fraction	$Y_{O_2,k}$	Input	-	0.04	0.10	0.07	0.03
13	Kiln dust %	$\dot{m}_{kd}$	Input	%	4.00	54.00	28.59	15.95
14	Pre-calcined meal temperature	$T_{o,pcm}$	Input	K	1055	1117	1090.04	30.74
15	O <sub>2</sub> molar fraction of outlet of precalciner	$Y_{o,O_2}$	Input	-	0.03	0.05	0.04	0.01
1	Apparent degree of calcination	$ADOC = \eta_{DOC,Apparent}$	Output	%	19.06	100	73.29	17.47
2	CO <sub>2</sub> molar fraction at precalciner outlet	$Y_{o,CO_2}$	Output	-	0.19	0.32	0.26	0.02
3	H <sub>2</sub> O molar fraction at precalciner outlet	$Y_{o,H_2O}$	Output	-	0.05	0.06	0.05	0.003

### 3 Results and Discussion

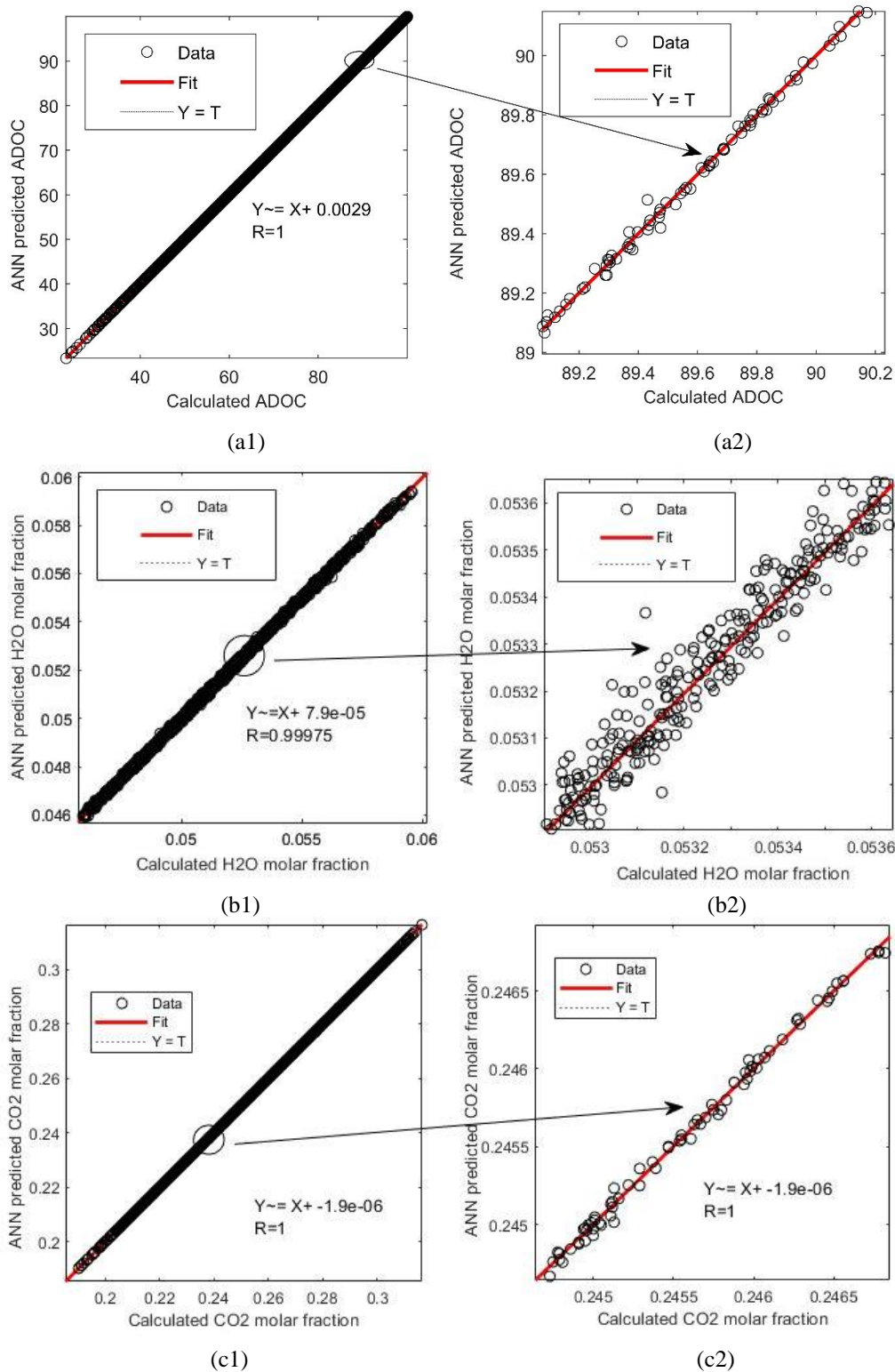
As mentioned earlier in Table 1, there were 5 regression model categories which were trained from the dataset. For each category, the model which gave the minimum RMSE was selected to predict the three output variables of apparent calcination degree, CO<sub>2</sub> molar fraction and H<sub>2</sub>O molar fraction. The summary of their statistical performance is shown in Table 3. It also shows the ANN model results. In addition, it also shows the *linear regression – classical* model result to give an understanding how the classical linear regression method deviates to other methods.

For predicting the apparent calcination degree, ANN gives the best results while GPR - rational quadratic method also shows successful results. Both Ensemble Bagged Tree and classical linear regression method show poor prediction results. ANN model also showed best performance for building relationship with inputs and CO<sub>2</sub> molar fraction and H<sub>2</sub>O molar fraction. Linear regression – stepwise algorithm was also successful for all the three outputs, but it demanded a considerable computational time compared to SVM and regression trees. SVM-medium gaussian and regression tree – fine, gave the third best results for CO<sub>2</sub> molar fraction prediction and H<sub>2</sub>O molar fraction prediction respectively. Training by Gaussian process regression algorithms were stopped due to high computational time for the CO<sub>2</sub> and H<sub>2</sub>O models.

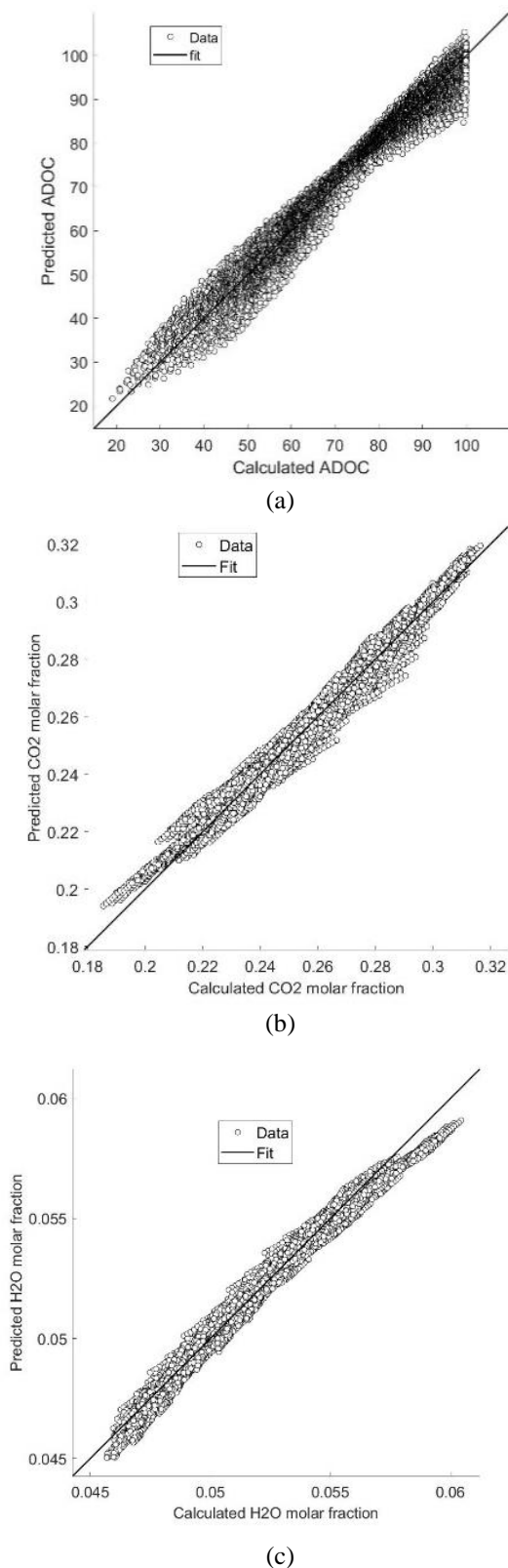
Figures 4 and 5 illustrate the prediction results for two different models. Figure 4 shows the performance by the ANN model compared to the MEB-based simulated data. ANN model predictions show a fit with R<sup>2</sup>=1, and due to the large number of samples tested, data scattering along the 1:1 line is not visible. Therefore, a small section of the x and y-axis was magnified to show the data swarm around the fitting curve. Results show that the ANN model can effectively formulate the relationship between these 15 parameters to the three output properties selected in this study. Figure 5 shows the comparison of *linear regression-classical* model compared to its MEB-simulated data. It shows the least performance compared to other algorithms for this dataset.

**Table 3.** Statistical details of the model performance

Parameter	Model name	RMSE	R <sup>2</sup>	MAE	Training time (sec)
Calcination degree	Linear regression – classical	3.224	0.97	2.5753	6
	Linear regression - stepwise	0.72151	1	0.57324	924
	Tree - Fine	1.9622	0.99	1.6375	8.5
	SVM - quadratic	0.91321	1	0.76053	15
	Ensemble – Bagged tree	2.4634	0.98	1.9776	36
	GPR – rational quadratic	0.03759	1.00	0.02804	1033
	ANN	0.019552493	1	-	-
CO <sub>2</sub> molar fraction (dry basis)	Linear regression - classical	3.3337 x 10 <sup>-3</sup>	0.98	2.5965 x10 <sup>-3</sup>	7
	Linear regression - stepwise	0.4666 x 10 <sup>-3</sup>	1	0.3633 x x10 <sup>-3</sup>	1314
	Tree - Fine	1.4738 x 10 <sup>-3</sup>	1	1.201 x10 <sup>-3</sup>	10
	SVM - medium gaussian	1.101 x 10 <sup>-3</sup>	1	0.898x10 <sup>-3</sup>	16
	Ensemble – Bagged tree	3.12 x 10 <sup>-3</sup>	0.98	2.471x10 <sup>-3</sup>	24
	GPR	Model training was stopped due to higher computation time			
	ANN	0.04157 x 10 <sup>-3</sup>	1	-	-
H <sub>2</sub> O molar fraction	Linear regression - classical	3.476 x 10 <sup>-4</sup>	0.98	2.085 x10 <sup>-4</sup>	3
	Linear regression - stepwise	0.711 x10 <sup>-4</sup>	1	0.5873 x10 <sup>-4</sup>	1218
	Tree - Fine	1.4129 x 10 <sup>-4</sup>	1	1.15x10 <sup>-4</sup>	5
	SVM - quadratic	1.6035 x 10 <sup>-4</sup>	1	1.378 x10 <sup>-4</sup>	6
	Ensemble – Boosted tree	3.647 x 10 <sup>-4</sup>	0.98	2.8007 x10 <sup>-4</sup>	12
	GPR	Model training was stopped due to higher computation time			
	ANN	5.97419 x 10 <sup>-5</sup>	0.99975	-	-



**Figure 4.** Performance by ANN model. Figure shows the comparison between the calculated data (MEB – based simulated data) and predicted data (by ANN model); (a1). ADOC; (b1) H<sub>2</sub>O molar fraction; (c1). CO<sub>2</sub> molar fraction; a2, b2, and c2 plots are magnified sections from their corresponding left sided plots



**Figure 5.** Performance by classical linear regression model. Figure shows the comparison between the calculated data (MEB – based simulated data) and predicted data (by classical linear regression model); (a). ADOC; (b) H<sub>2</sub>O molar fraction and (c). CO<sub>2</sub> molar fraction

There are both advantages and disadvantages between different machine learning algorithms. Their performance also heavily depends on the type of data. The theory behind these algorithms are not mentioned in this paper and can be found in literature such as in (Shwartz and David, 2014). In general, it is said that SVM and regression trees have fast prediction and training speeds, but they are suitable for handling minor problems and prone to overfitting (Bonaccorso, 2017). Ensemble gives high accuracy and performance for small and medium-size datasets, but tuning is required. Gaussian process is an effective algorithm for both regression and classification. A Gaussian process is a probability distribution over possible functions, and can deal effectively with data uncertainty (Irwin, 1997). The most critical drawback of GP regression is higher computation time. In this study, using GPR algorithm were terminated for H<sub>2</sub>O and CO<sub>2</sub> molar fraction prediction. The advantage of modelling by ANN is that the model can be established directly with the input and output data of the application when there is less prior knowledge of the application. It is suitable for the highly nonlinear and uncertain system. ANN model has better online correction capability (Abiodun, Jantan et al., 2018). But it uses a large memory, and training speed can be slow. However, the choices of the quantity and quality of training data, learning algorithm, and topology and type of the network are all critical to the performance of a soft sensor model.

The data used to train the models in this study are simulated data. Some data points used in model development may not be practical in plant operation. The process data generated from real-time plant operation is a mixture of noise from raw materials, energy inputs, equipment, system running state, and time-varying chemical and physical parameters of raw materials and products. Therefore, if the real process data can train the models mentioned in this study, it will be an excellent opportunity to assess the results of this feasibility study. Plant operators can use such models tuned for a prolonged period to reduce downtime and take decisions before the results of offline lab samples arrive. Since plant data always include noise and undefined variations, the way they fit to the models might be different than reported in this study. Therefore it is always recommended to tune the model with a number of plant data before the models are used directly for practical applications.

#### 4 Conclusion and Recommendation

This work describes a regression attempt to determine the apparent degree of calcination, CO<sub>2</sub> molar fraction and H<sub>2</sub>O molar fraction in a cement precalciner system by formulating relationships between several input variables. Different types of machine learning algorithms were tested to test their suitability to build relationships with the input data and output data. Since



this is a feasibility study, synthetic data were used to train the models. These data were obtained from applying mass and energy balance. Results show that a number of machine learning algorithms show good performance with respect to the classical linear regression method. Several factors affect the calcination process in the precalciner and adding their contribution as model inputs to tune the developed models are recommended to increase the model robustness. In addition, it will also provide clues whether sampling frequency should be increased or not for experimentally measured parameters used to calculate the precalciner performance.

In particular, training a model using synthetic data can be viewed as a learning process. The advantage of using synthetic data from such theoretical models is that the number of data points can be increased to decrease the error inexpensively. The results can be viewed as a guide for a proposal distribution generator for approximate inference and can be used to draw a formal connection between inputs to optimize network parameters. As the second step of this study, testing the models by selected process data that represent extreme and typical plant operation conditions is recommended. This will lead to develop more realistic models based on the actual plant data. The system boundary used for the model was the precalciner system. However, the model can be more meaningful if the system boundary can be expanded to cover the entire pyroprocessing unit.

## Acknowledgements

Financial support from University of South-Eastern Norway, Telemark Fylkeskommune and the Norwegian Research Council is greatly acknowledged.

## References

- O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad. State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11), e00938-e00938, 2018. doi:10.1016/j.heliyon.2018.e00938
- G. Bonaccorso. *Machine Learning Algorithms*: Packt Publishing Ltd., 2017.
- IKN GmbH. CO2 capture from cement production; D 8.1 Status Report on Calciner Technology Revision 2, 2016.
- Y. Z. Gang, and L. Hui. Soft sensor for apparent degree of calcination in NSP cement production line. Paper presented at the 2nd *International Conference on Computer and Automation Engineering (ICCAE)*, 2010
- M. K. Griparis, F. N. Koumboulis, N. S. Machos, and I. Marinou. Precalcination in cement plants (system description and control trends). *IFAC Proceedings Volumes*, 33(20), 273-278. 2000. doi:10.1016/S1474-6670(17)38062-X
- IEA. CO2 capture in the cement industry; Technical Study Report number 2008/3. 2018.
- J. D. Irwin *The Industrial Electronics Handbook (The Electrical Engineering Handbook Series)* CRC Press, 1997.
- Mathworks. Regression Learner. Retrieved from <https://www.mathworks.com/help/stats/regressionlearner-app.html>, 2021.
- H. Mikulčić, E. von Berg, M. Vujanović, P. Priesching, L. Perković, R. Tatschl, and N. Duić. Numerical modelling of calcination reaction mechanism for cement production. *Chemical Engineering Science*, 69(1), 607-615. 2012. doi:10.1016/j.ces.2011.11.024
- N. Mohammadhadi. *Multi-phase flow and fuel conversion in cement calciner*. PhD thesis. Kgs. Lyngby: Technical University of Denmark (DTU), 2018.
- J. Osmic, E. Omerdic, E. Imsirovic, T. Smajlovic, and E. Omerdic. Identification and Control of Precalciner in the Cement Plant. In: J.A.Gonçalves, M.Braz-César, and J.P.Coelho (eds) *CONTROLO 2020*. Lecture Notes in Electrical Engineering, vol 695. 2020. Springer, Cham. doi: 10.1007/978-3-030-58653-9\_12
- S. S.-Shwartz and S. B.-David. *Understanding Machine Learning: From Theory To Algorithms*: Cambridge University Press. 2014.
- L.-A. Tokheim. *The impact of staged combustion on the operation of a precalciner cement kiln*. PhD thesis. NTNU, 1999.
- WWFI. A blueprint for a climate friendly cement industry; How to Turn Around the Trend of Cement Related Emissions in the Developing World. 2008.
- B. Yang, H. Lu, and L. Chen. BPNN and RBFNN based modeling analysis and comparison for cement calcination process. Paper presented at the *Third International Workshop on Advanced Computational Intelligence*, 2010.