Simulation-Assisted AI Modeling for Glass Quality Prediction

A. Ghosh, L. Thielen, J.P. van der Dennen
(CelSian Glass & Solar B.V., The Netherlands).

Abstract

The production of high-quality glass demands precise control over furnace parameters, raw material inputs, and a thorough assessment of the final glass product. Quality and defect levels are crucial factors influencing efficiency, cost, waste, and sustainability in glass manufacturing. Given the multitude of furnace parameters, often of the order 100, determining quality through numerical simulation alone remains a challenging problem. Adding to this complexity is the time delay between the formation of molten glass and the detection of defects in the final product. To achieve good quality glass, the residence time of the glass in the furnace needs to be at least 8-12 hours. This residence time is not constant and highly depends on the process parameters. Changes made to the inputs will affect the quality of the final product at a much later stage. An application that predicts the upcoming glass quality as a function of current and previous inputs supports optimal furnace performance. CelSian addresses these challenges by integrating machine learning with furnace simulation to predict glass quality based on process parameters and user input variables. Celfos, powered by CelSian's CFD simulation package GTM-X, simulates furnace dynamics from user inputs, capturing changes in parametric values over time. By training an AI model on a large dataset, Celfos predicts defect counts over time. This approach combines the physics-driven simulation of GTM-X with the predictive power of AI, offering a novel pathway for proactive quality control in glass production. An important aspect of the development is a parallel, governmental-funded, project to speed up the CFD code significantly. This is done through AI-enforced solvers and usage of GPU's. The currently ongoing research projects for glass quality prediction show promising results. This article shows some of the results achieved and a forecast for future improvements. Also, the implementation of AI to speed up the CFD is briefly discussed.

1. Introduction

The art of glass production has a long history and is entwined with the progress of human civilization. Although glass production has been studied over millennia, difficulties are still faced at an industrial level. Glass production is governed by a multitude of parameters ranging from furnace control parameters, raw materials and through the assessment of the final product for defects. Precise control of each parameter is necessary for the production of defects-free glass. Quality and number of defects are crucial factors which influence efficiency, cost, waste, and sustainability in the process of glass manufacturing. Over time multiple studies have been performed to numerically simulate the glass furnace [1-5]. These numerical methods use computational fluid dynamics (CFD) techniques to simulate all relevant phenomena in the complete glass melting process such as: flow, temperature, melting, multi-phase flow, electrical fields, chemical reactions, radiation, and combustion. However, one of the major drawbacks of these numerical simulations is the inability to predict the quality of glass as the end of the production process.

Industrial furnaces are large and hence multitude of parameters, often together, control the quality of the final glass product. This also includes the quality of the raw materials being used which changes over time due to changes in supply. The furnace itself goes through repairs and parts are replaced due to aging. Many unforeseen parameters play a role in the quality of the final glass product. Hence it is non-trivial to predict the quality of the glass through numerical simulations. As a result, this is an active field of research at both academic and industrial institutions.

CelSian has developed a numerical model powered by machine learning to predict the quality of glass from the different furnace parameters. The model is trained on a real-life dataset obtained from our industry partners. By combining the furnace dynamics simulations from CFD [4,5,6], the model uses machine learning to predict the number of defects in the produced glass. The residence time of the molten glass inside the furnace is modeled using tracers on the CFD solution which is used to pre-process the furnace parameter data. The furnace parameters include temperatures from different thermocouples at different parts of the furnace including melter and the crown, speed, tonnage, fuel flow rate, etc. The full process is shown in Figure 1.

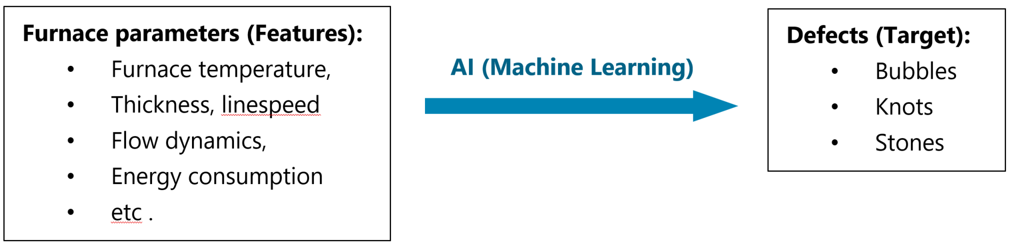

Figure 1: Illustration of the problem statement

This paper is structured in the following way: section 2 describes the model of the numerical method, section 3 elaborates the results obtained from the model and section 4 concludes the current article with a brief discussion about the future prospects.

2. Model Description

One of the governing factors of the quality of the final glass product is the dynamics of the molten glass inside the furnace. Complex physical phenomena involving fluid dynamics, chemical reactions, heat transfer, combustion, etc. occur inside the furnace. Figure 2 represents a schematic of a glass-melting furnace. The raw materials are pushed inside the furnace which contains molten glass. Gaseous fuel is injected into the upper half of the furnace where combustion takes place. The heat is radiated from the combustion space of the furnace which heats the raw materials and it ultimately melts. Convective flow occurs inside the molten glass that enhances mixing and proper heat distribution which aids the melting of the raw materials.

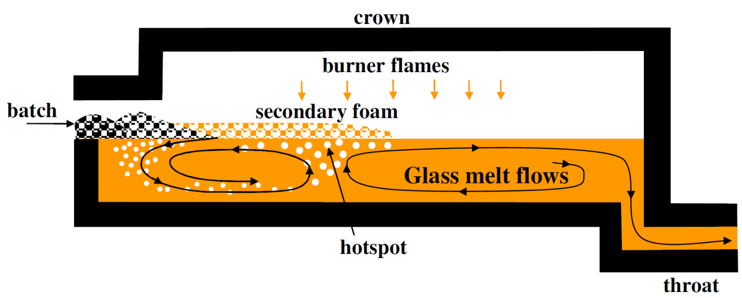

Figure 2: Schematic view of a glass melting furnace [4].

The molten glass is then processed into the required final product (container glass, float glass, fibre glass, etc). Accurate modeling of the complex phenomenon inside the glass furnace is necessary for the prediction of the final glass quality.

The main work described in this paper is to create a machine learning model to correlate the number of defects of the final glass product with the different furnace parameters, which are obtained from different sensors places in different parts of the furnace. One of the main governing factors on the quality of glass is the temperature of the molten glass at different parts of the melter (the bottom half of the furnace that holds the molten glass). The temperatures are detected using thermocouples.

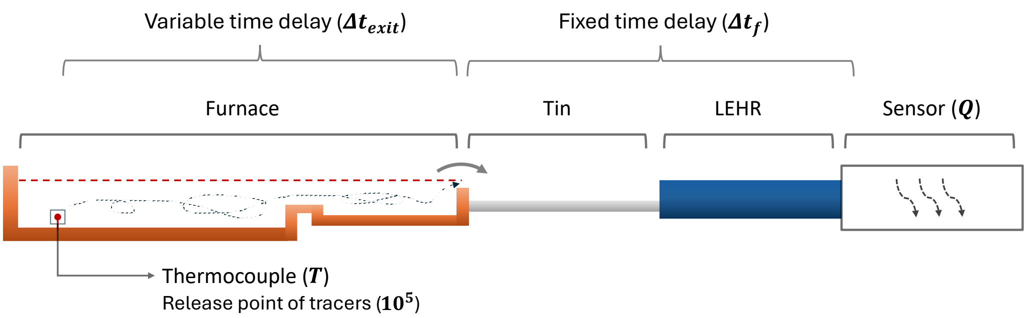

Figure 3: Schematic of float glass production process. Visualization of a tracer from start to exit of the line.

Due to the complex geometry and flow profile in the melter, there is a delay in the temperature $T(t)$ measured at time t and the quality metric, say the number of defects, detected at time t, say $Q(t)$. This delay comes from mainly two factors, the glass is passing through the lehr and tin bath after being melted before reaching the quality sensor, so the delay depends on the line-speed of the production process and can be overcome by using a fixed delay in time, Δt_f. The second delay is more complex to deal with as it originates from the complex fluid flow of the molten glass inside the melter. The unit volume of molten glass, the temperature of which a thermocouple detects at time t, does not instantly go to the exit of the furnace, but rather follows a complex path. Also, the unit volume of molten glass breaks up along the path to the exit. This delay can be simulated using tracers as shown in the schematic of the Figure 3.

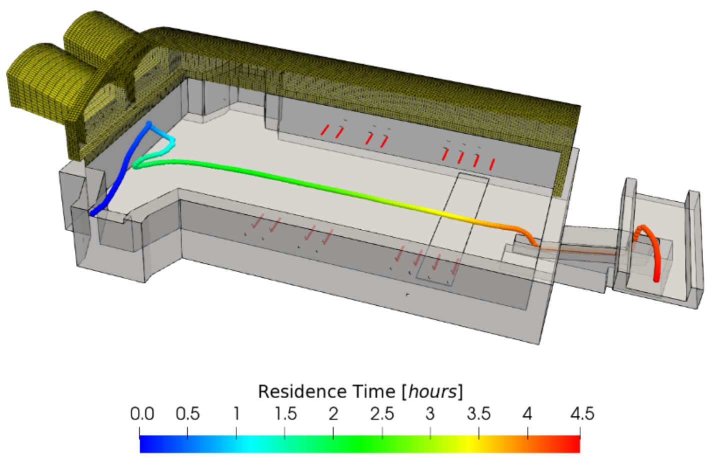

Figure 4: Trajectory of the tracer particles. This simulation is a representation of tracer particle study in a generic glass furnace.

The simulations of the glass furnace are performed on our in-house solver GTM-X which is a boundary-fitted CFD-model that simulates the complete glass melting process. The flow equations are solved using a finite-volume method on a body-fitted, multi-block, structure grid [10]. The solver is parallelized using MPI and the blocks are distributed over a number of processors which sufficiently shorten the computational time to within the order of 1 to 2 days.

The simulation of tracers is performed with a steady-state solution of the CFD model. As shown in Figure 4, for each thermocouple location, 60,000 inert tracers are initialized from a volume of say 100mm sided cube. The volume of fluid is assumed to influence the said thermocouple. The particles released follow the path of the streamlines. Since each streamline is different, the tracers take different paths and hence amounts of time to reach to the exit of the furnace. This can be plotted in the form of a residence time distribution plots as shown in Figure 5.

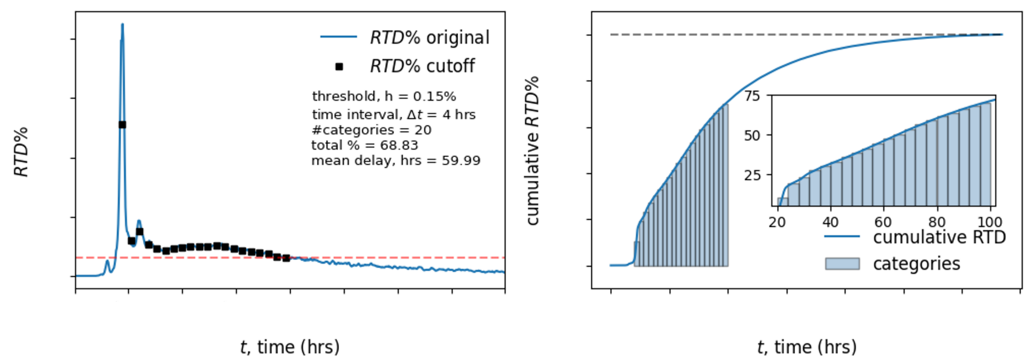

Figure 5: A typical residence time distribution (RTD%) of a generic glass furnace. The red dashed line denotes the threshold value, above which the data is considered. The blue shaded area of the right plot denotes the cumulative tracers considered for the data.

Figure 5 shows the residence time distribution (RTD) of a tracer study from a single thermocouple. A minimum threshold h% is set to focus on the tracers that stay in the furnace for a longer period of time. The RTD is then discretized into a set interval and the residence times for each discretized section is documented. These residence times, represented as Δt_{exit}, along with the fixed delay (Δt_f) will be used in the pre-processing of the data for the training of the machine learning model.

Machine learning model

The goal is to correlate the number of defects of the final glass product with the values of the different furnace parameters. The furnace being a highly complex instrument is governed by multiple parameters, which is of the order 100. These parameters have different levels of dominance over the final glass quality. Hence machine learning is one of the best solutions for correlating between multiple input features to a single target parameter.

The data used to train the machine learning model is collected in collaboration with a float glass manufacturer. Raw sensor and thermocouple data from the float glass furnace is uploaded to their servers and subsequently transferred to the Snowflake data storage facility. The raw data is accessed using SQL commands, performing preprocessing, and ultimately using the cleaned dataset to train machine learning models. Depending on the type of sensor, the data stored on Snowflake are around one minute interval and total 2 years' worth of data is stored and updated in real time. One year's worth of data from a furnace was used. The data, especially the thermocouple temperature data, needs to be pre-processed before being used in the training process. The delay between the reading of the thermocouple temperature and the detection of defects is obtained using the residence time distribution from the CFD simulations. From the discretized RTD, the time required for the glass tracer to reach the exit of the furnace from the start thermocouple is categorized for each discrete segment of the RTD. Using these categorized time values and the historical temperature data for each thermocouple obtained from the plant furnace, different categories of temperature data are created per training to each delay category. The delay for each category, the time delay is denoted as Δt_{exit}^1, Δt_{exit}^2, …, Δt_{exit}^n , where the superscript {1, 2, …, n} denotes the category of the discretized RTD. Assuming the data is pre-processed with a period of Δt, the pre-processing of the temperature can be represented as shown in Table 1.

Table 1: Modification of temperature data to accommodate the delay due to residence time inside the furnace and also the fixed delay due to the lehr and tin bath. Q(t) represents the quality metric, which is number of defects in this article, measured at time, t.

Time	T1	T2	…	Tn	Q
$\boldsymbol{t}$	$T_1(t - \Delta t_f - \Delta t_{exit}^1)$	$T_2(t - \Delta t_f - \Delta t_{exit}^2)$	$\cdots$	$T_n(t - \Delta t_f - \Delta t_{exit}^n)$	$Q(t)$
$\boldsymbol{t + \Delta t}$	$T_1(t - \Delta t_f - \Delta t_{exit}^1 + \Delta t)$	$T_2(t - \Delta_f - \Delta t_{exit}^2 + \Delta t)$	$\cdots$	$T_n(t - \Delta t_f - \Delta + \Delta t)$	$Q(t + \Delta t)$

$\boldsymbol{t + 2\Delta t}$	$T_1(t - \Delta t_f - \Delta t_{exit}^2 + 2\Delta t)$	$T_2(t - \Delta t_f - \Delta t_{exit}^2 + 2\Delta t)$	$\cdots$	$T_n(t - \Delta t_f - \Delta t_{exit}^n + 2\Delta t)$	$Q(t + 2\Delta t)$
$\boldsymbol{\cdots}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\boldsymbol{t + N\Delta t}$	$T_1(t - \Delta t_f - \Delta t_{exit}^n + N\Delta t)$	$T_2(t - \Delta t_f - \Delta t_{exit}^2 + N\Delta t)$	$\cdots$	$T_n(t - \Delta t_f - \Delta t_{exit}^n + N\Delta t)$	$Q(t + N\Delta t)$

This preprocessing of temperature for all the thermocouples is performed based on the Table 1. Also including all other furnace parameters, the total input variables (input features) for the training of the machine learning model is around 1000.

For the machine learning process, a feed forward multi-layered perceptron (FFMLP) was used with 3-4 layers.

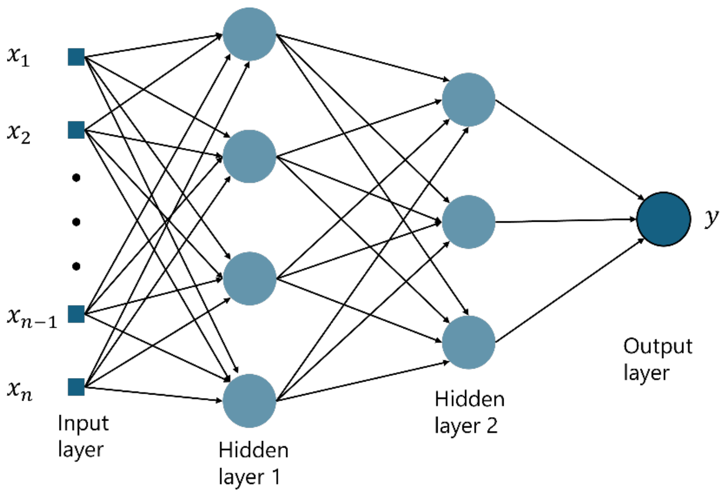

Figure 6: Schematic of a feed-forward multi-layered perceptron.

Figure 6 shows a schematic of a version of a feed-forward multi-layered perceptron (FFMLP). The preprocessing of the data and the model is trained using Python [7] and TensorFlow [8] respectively. A mean squared error (mse) function was used as the loss function. For updating the weights and the biases of the model during the training process, the Adams [9] optimizer was used. The metrics used for measuring the accuracy of the model are root mean squared error (RMSE) for the error and the coefficient of determination, also known as R^2 score. The relevant equations can be shown as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{N}^{i=1}\left(\widehat{y_i}^2 - y_i^2\right)} \quad (1)$$

$$R^2 score = 1 - \frac{\sum_{n}^{i=1}(y_i - \hat{y})^2}{\sum_{n}^{i=1}(y_i - \overline{y_i})^2} \quad (2)$$

where $\widehat{y_i}$ is the predicted data and y_i is the real (test) target data. The mean of the test target is given by $\overline{y}$. The accuracy can also be calculated from the RMSE error as follows:

$$accuracy = 1 - \frac{RMSE}{y_{max} - y_{min}} \quad (3)$$

where y_{max} and y_{min} are maximum and minimum values of the target in the test dataset. Multiple different learning rates over a range of 0.00001 to 0.001 are used to model so as to find the optimum value that reduced the noise, increased stability, and was trained the fastest.

3. Results and Discussion

In this section, the results of the trained model are shown using real-life data from the furnace. The data is obtained from industrial partners who provided us with the data to train the model.

This paper focusses on the defects which are the main cause of reject in glass manufacturing: bubbles. The features used in the model are temperature data from the melter and different furnace parameters like linespeed, tonnage, crown temperature, etc. Multiple features were tested to find the most optimum features that give the best correlation between the furnace parameters and the target defects.

Multiple tests were performed to determine the optimum hyperparameters including which type of optimizer to use and the formula for loss calculation, value of the batch size, learning rate, number of layers and densities of each layer, etc. Also, the tests were performed to find the combination of furnace parameters that correlate the highest with the number of defects.

The data used to run the training of the machine learning model is obtained through private communications with our industrial partners. Due to the proprietary nature of the data, the raw data cannot be shown publicly. The total dataset is for 2 years of data, containing around 30,000 datapoints which is split into training and testing dataset, with 20% going to the test dataset. The dataset contains raw data from the hundreds of sensors located at various parts of the furnace. The most relevant ones were chosen for the development of the machine learning model.

For the first test, we choose one thermocouple data and a few other furnace parameters. The temperature data from the thermocouple is pre-processed using the RTD as discussed in the previous section. Around 150 features is created with a single target (number of bubbles). A 2 layer model architecture was used for the test with layer density of {256, 64}.

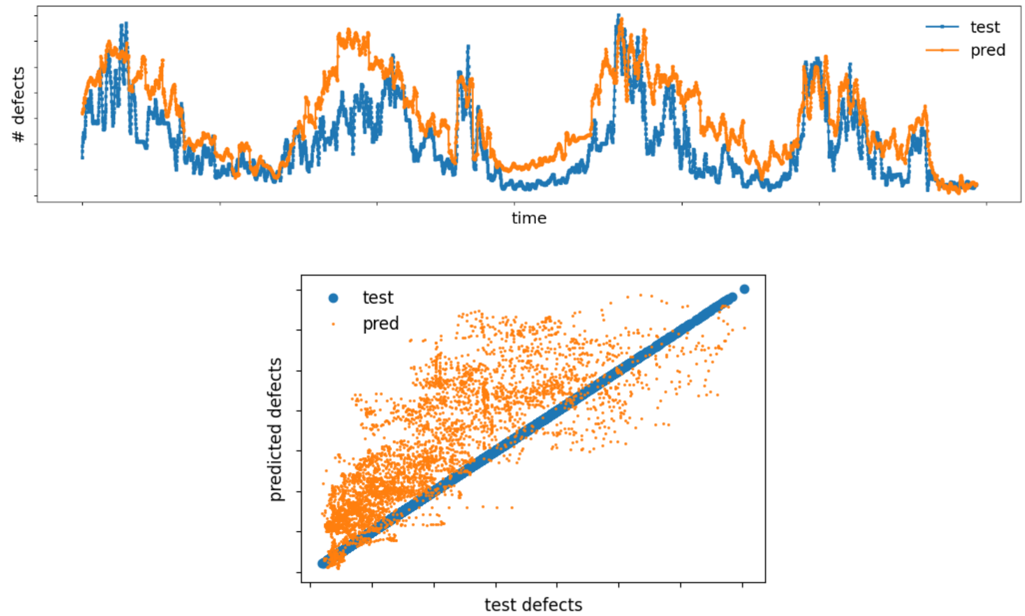

Figure 7: (a) Example of Predicted (orange) defects over a period of time compared to the actual defects (blue) defects detected during the same period. (b) Actual (test) defects vs predicted defects. The actual defects are in blue and the predicted defects are in orange. The predicted defects match closely with the actual defects with a slight over-prediction over the period.

Using the model described in the previous section, the model gives a prediction of the defects as shown in Figure 7. It shows a period of around 3 months of the data target data (number of defects) as compared to the predicted target for the same period. The predicted data matches quite well with the test data with a slight over-prediction. Since the model is designed to predict defects, at an industrial level, under-prediction of defects is more hazardous than over-prediction and hence slight over-prediction is preferred. The accuracy of this example is around 78%.

Another test is performed where 4 thermocouples are used as input parameters along with other furnace parameters. Similar to the previous test, the melter temperature data are pre-processed using the RTD method as described in the previous section which incorporates the delays over time in a single dataset. Due to the high number of temperature data, the RTD creates features almost 1000 in number. Due to the higher number of input features, the model used here has 3 layers with structure {728, 256, 128}. The model is run for 1000 epochs at which point the validation loss and becomes constant even though the training loss decreases. To avoid over fitting, the model training is stopped around 1000 epochs.

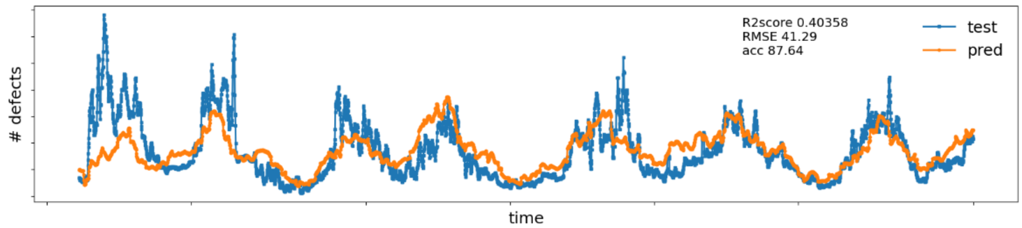

Figure 8: Example of Predicted (orange) defects over a period of time compared to the actual defects (blue) defects detected during the same period.

Figure 8 shows the comparison between the test target data in blue and the predicted target data in orange. The model gives a good accuracy with the metrics R^2 score=0.40358, RMSE=41.29, accuracy=87.64%. The model is accurate enough to provide real time prediction of defects from furnace parameters. Comparing between the results of Figure 7 and 8, it can be observed that the model has more accurate predictive capabilities with higher number of furnace parameters.

4. Conclusion and future work

In this article a physics assisted machine learning model was introduced for the purpose of correlating glass furnace parameters with number of defects in the final glass product.

The correlation between the furnace parameters and the glass defects are non-trivial and hence GTM-X is used to obtain CFD simulations of a glass furnace. These simulations provide us with the dynamics of molten glass flow inside the furnace. From these simulations the delay between the glass being present at a furnace sensor (thermocouple) and the final glass product under the defect-scanner is calculated using residence time distributions (RTD) using tracers. The time delays are obtained from the discretized RTD which are used to create temperature data with different delays.

The raw data from the furnace is pre-processed to remove any outliers. The model is trained with a complex feed forward multi-layered perceptron (FFMLP). Multiple tests are performed to optimize the hyperparameters to increase accuracy and the learning speed of the model. Also, tests are performed to find the combination of furnace parameters that correlate the highest with the defects. Ultimately an accuracy of 87% with R^2 score of 0.4 is obtained with the test dataset.

Future plans for further development are in the pipeline. More complex models including convolutional models and memory-based neural network are under development to further improve the temporal prediction capabilities.

5. References

[1]. A. Abbassi and Kh. Khoshmanesh, “Numerical simulation and experimental analysis of an industrial glass melting furnace,” *Applied Thermal Engineering*, vol. 28, no. 5–6, pp. 450–459, Apr. 2008, doi: https://doi.org/10.1016/j.applthermaleng.2007.05.011

[2]. J. Raič *et al.*, “CFD simulation aided glass quality and energy efficiency analysis of an oxy-fuel glass melting furnace with electric boosting,” *Energy Conversion and Management: X*, vol. 15, p. 100252, Jun. 2022, doi: https://doi.org/10.1016/j.ecmx.2022.100252

[3]. O. Verheijen, Andries Habraken, and H. Gramberg, “Modeling of Heat Transfer and Gas Flows in Glass Furnace Regenerators,” *Ceramic engineering and science proceedings*, pp. 201–206, Apr. 2015, doi: https://doi.org/10.1002/9781119117490.ch17

[4]. L. Thielen, A.M. Lankhorst, B.D. Paarhuis and O.S. Verheijen, “NAFEMS Simulation of Multi-Physical Phenomena in Glass Melting Furnaces,” *NAFEMS*, 2025.

[5]. A. Faber, Mathi Rongen, Adriaan Lankhorst, and Domingos De Sousa Meneses, “Characterization of high temperature optical spectra of glass melts and modeling of thermal radiation conductivity,” *International Journal of Applied Glass Science*, vol. 11, no. 3, pp. 442–462, Feb. 2020, doi: https://doi.org/10.1111/ijag.15111.

[6]. Anne Jans Faber, Mathi Rongen, and Domingos De Sousa Meneses, “High-temperature near-IR spectral properties and thermal radiation conductivity of (un)colored silicate glass melts,” *International Journal of Applied Glass Science*, vol. 14, no. 2, pp. 189–200, Aug. 2022, doi: https://doi.org/10.1111/ijag.16603.

[7]. G. Van Rossum, & F. L. Drake, (2009). “Python 3 Reference Manual,” *Scotts Valley, CA: CreateSpace*.

[8]. TensorFlow Developers, “TensorFlow,” Zenodo, Oct. 2024, doi: https://doi.org/10.5281/zenodo.13989084.

[9]. D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” arXiv.org, Dec. 22, 2014. https://arxiv.org/abs/1412.6980.

[10]. J. H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*. 1996.