Real-time prediction of interstitial oxygen concentration in Czochralski silicon using machine learning

Kentaro Kutsukake\textsuperscript{1}, Yuta Nagai\textsuperscript{2}, Tomoyuki Horikawa\textsuperscript{2}, and Hironori Banba\textsuperscript{2}

\textsuperscript{1}Center for Advanced Intelligence Project, RIKEN, Nhonbashi, Chuo-ku, Tokyo 103-0027, Japan

\textsuperscript{2}GlobalWafers Japan Co., Ltd., Seiro, Niigata 957-0197, Japan

E-mail: kentaro.kutsukake@riken.jp

We developed a machine learning model to predict interstitial oxygen (Oi) concentration in a Czochralski-grown silicon crystal. A highly accurate prediction can be ensured by selecting the appropriate experimental parameters that represent the change in the furnace conditions. A neural network was trained using the dataset of 450 ingots, and its prediction error for the testing dataset was $4.2 \times 10^{-16}$ atoms cm$^{-3}$. Finally, a real-time prediction system was developed wherein the crystal growth data are input into the model, and the Oi concentration at the current growth interface is calculated immediately. © 2020 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd

Material evaluation is a fundamental process in the production and development of materials; it is accompanied by a time lag between the initiation of fabrication and the completion of evaluation. This poses a significant problem in the process of evaluating the bulk crystalline materials that undergo time-consuming preparation processes wherein the samples are cut and sliced. The evaluation of the interstitial oxygen (Oi) concentration\textsuperscript{1}–\textsuperscript{3} in Czochralski (CZ) silicon crystals typically requires four days: two days for crystal growth, one day for grinding and cutting, and one day for measurements. Furthermore, there are few discrete positions at which the Oi concentrations are measured in a crystal ingot. It is important that the process of material evaluation be continuous and have no time lag between the different stages of the process.

Predicting impurity concentration using crystal growth parameters is one of the ways to solve this problem. Numerical analysis and crystal growth simulation help predict the impurity distribution in a crystal ingot by modeling the transfer of impurities during the crystal growth. Many studies have reported the mechanism of oxygen transport in a furnace during the growth of silicon crystals.\textsuperscript{5}–\textsuperscript{8} In the CZ silicon crystal growth, the melt is contained in a silica (quartz) crucible; as the process progresses, the walls of the crucible dissolve into the melt and therefore oxygen is transported inside the silicon melt by diffusion and convection of the melt. Generally, this is how oxygen impurity is incorporated into the silicon melt during the CZ silicon crystal growth.\textsuperscript{9}–\textsuperscript{13} Although most of the oxygen evaporates from the surface of the melt as SiO gas, some oxygen is incorporated into the silicon crystal. On the basis of these oxygen transport mechanisms, two-dimensional and three-dimensional numerical models of oxygen transport have been developed to analyze the relationship between oxygen concentration in CZ silicon crystals and the various control parameters during crystal growth.\textsuperscript{14}–\textsuperscript{19} Brown et al. qualitatively predicted the Oi concentration in the crystal by analyzing the heat transfer and gas flow in a CZ furnace.\textsuperscript{20}–\textsuperscript{22} Togawa et al. used numerical analysis to investigate the impact of crucible rotation rate, an essential parameter that influences melt convection and Oi concentration in CZ silicon crystals.\textsuperscript{23,24}

However, it is difficult to accurately model the actual flow pattern and oxygen transport inside the crucible even using a three-dimensional simulation. This is because there are many parameters in an actual crystal growth process that significantly affect the impurity concentration which are difficult to formulate. In addition to the process parameters, factors such as age of the graphite crucible and size of the ingot shoulder that vary slightly with each growth cycle also influence Oi concentration. These are difficult to incorporate in the simulation models. Furthermore, simulation is time-intensive which poses a significant problem during investigation of the effects of various parameters on Oi concentration and the interactions between the parameters.

Non-parametric machine learning models such as neural networks can incorporate the influence of experimental parameters that are difficult to formulate. Thus far, most of the machine learning models of crystal growth have been simulation-based that have systematically produced data using numerical simulations without noise or fluctuation.\textsuperscript{25}–\textsuperscript{29} In addition to this noise-related problem, for the machine learning modeling of practical crystal growth, the number of parameters that should be considered is higher than that of simulation results, as mentioned above. In this study, we trained a machine-learning model to predict Oi concentration in a CZ silicon ingot using a dataset of experimental data (consisting of 100 parameters) from 450 ingots grown in the same furnace in the presence of a magnetic field. The crystal diameter was 300 mm. The Oi concentration range of the samples was between $0.7 \times 10^{18}$ and $1.7 \times 10^{18}$ atoms cm$^{-3}$, and the resistivity range was between 1.5 and 78.1 $\Omega$ cm. The Oi concentration was measured via Fourier transform infrared spectroscopy (FT-IR) of wafer samples cut from several positions of each ingot. Figure 1 shows a schematic of the data structure for the machine learning model. The experimental data that is input to the machine learning model can be categorized into three groups: fixed parameters such as age of the graphite heater and crucible, process parameters such as ingot pulling rate, crucible rotation rate and crystal rotation rate, and monitored parameters such as temperature and ingot diameter. The process parameters are controllable and set by the operator, and the monitored parameters are only for observation and

Content from this work may be used under the terms of the Creative Commons Attribution 4.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. © 2020 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd
not directly controlled. In the case of Oi prediction using the crystal growth simulation, the monitored parameters are not input parameter because the calculation results of monitored parameters are determined by the process parameter values. However, in the case of practical crystal growth, though the process parameter values remain the same, the monitored parameters vary in each ingot owing to the differences in the variation in furnace conditions. Thus, we added the monitored parameters to the input dataset of the machine learning model in order to contain information about the furnace condition. We also directly take the influence of the furnace condition into account by adding the fixed parameters that change by each ingot. The process and monitored parameters are time series data with more than 20,000 values for an ingot. Thus, the total data size is 100 (number of parameters) × 20,000 (time direction) × 450 (number of ingots) = 900 million. From this huge data set of the acquired parameter values, we extracted data rows corresponding to the ingot length position where Oi concentration was measured. Merging the extracted rows of the process and monitored data with fixed data and Oi concentration values, we prepared the dataset for the machine learning. Through this process, there were as many number of data rows as there were Oi data, which was 2209 in this study. This is the actual data-set (number of rows) used against the effective number of variables in the machine learning model. Subsequently, this dataset was categorized into 1554 training data and 655 test data; particularly, each ingot datum was assigned to one of these data groups to prevent data contamination. Finally, the parameter values were standardized using the standardization formula: \( x' = (x - u)/s \), where \( x' \) and \( x \) are the standardized and original parameter values, and \( u \) and \( s \) are the mean and standard deviation of the parameter values of the training dataset, respectively.

Before feeding the dataset into the machine learning model, we performed correlation analysis on the input parameters. Some parameters showed a strong linear correlation; therefore, one parameter from every pair that had a correlation coefficient higher than 0.90 was excluded from the input of the machine learning model. For example, the correlation coefficient between the process time and ingot length or between the set values and actual values of the pulling rate was more than 0.90; hence, one parameter from each of these pairs was removed from the input. This process reduced the number of parameters to 43.

The fully-connected feedforward neural network model was made and trained using the Keras library in Python.30) The number of layers, nodes in each layer, and epochs were 3, 47, and 20,000, respectively. Dropout and batch normalization were used after every fully connected layer. A sigmoid activation function was applied, and the loss was measured by mean squared error with layer weight regularizer. To optimize all the learning parameters in the network, backpropagation based on the Adam scheme31) was implemented. The above-mentioned hyperparameter values were optimized by a combination of manual exploration and Bayesian optimization.

Figure 2 shows the relationship between the measured and predicted Oi concentrations for the training and test dataset. The coefficient of determination (R² score), mean absolute error (MAE), and root mean squared error (RMSE) for the test data were 0.94, 2.7 × 10¹⁶ atoms cm⁻³, and 4.2 × 10¹⁶ atoms cm⁻³. This prediction error is low enough for the practical estimation and control of Oi concentration. Note

![Fig. 1. (Color online) Schematic of data structure for the machine learning model. CR: crucible rotation rate. CA: graphite crucible age (number of usage). The experimental data was the time series data of the progress of crystal growth with respect to the growth length. The machine learning data set was constructed by extracting and combining the huge data set consisting of process, monitored, and fixed parameter data.](image)

that the minimum value of the predicted Oi concentration sharply locates at $0.9 \times 10^{16}$ atoms cm$^{-3}$. This is the result of the learning to balance between prediction accuracy and overfitting. There is a possibility that prediction accuracy around the lower limit is improved by further optimization of the neural network structure and hyper parameters.

We investigated the effects of the monitored and fixed parameters on the prediction accuracy of the machine learning model. The 43 input parameters of the above-mentioned neural network model consisted of 21 fixed, 7 process, and 15 monitored parameters. We also made neural network models using only fixed and process parameters (without monitored parameters), monitored and process parameters (without fixed parameters), and process parameters (without monitored and fixed parameters). RMSEs of the four models for the test dataset are shown in Fig. 3. The RMSE of the model trained with all the parameters is measurably smaller than that of the other models; $4.8 \times 10^{16}$, $4.8 \times 10^{16}$, and $5.4 \times 10^{16}$ atoms cm$^{-3}$ for the models trained without monitored parameters, without fixed parameters, and without monitored and fixed parameters, respectively. This suggests that the effect of monitored and fixed parameters on the prediction accuracy is significant. In addition, the variation of the RMSE is remarkably large in the model trained without fixed parameters. This indicates that the fixed parameters are also effective to generalization of the model. The process parameters were controlled to keep the ingot diameter constant during the growth of the ingot body; further, the regulation of the diameter determines the Oi concentration. The Oi concentration is also influenced by the furnace conditions. The monitored parameters reflect the time changes in the furnace conditions during the growth, such as SiO deposition condition and the melt surface positions with respect to the crucible. The fixed parameters represent the difference of the furnace conditions between the ingot growth, such as physical properties of the graphite heaters and crucibles. Therefore, they have a significant effect on the accuracy and generalization of the prediction.

The prediction accuracy achieved by the machine learning model, indicated by $\text{MAE} = 2.7 \times 10^{16}$ atoms cm$^{-3}$, is still greater than $3 \times 10^{15}$ atoms cm$^{-3}$, which shows the repeatability of the FT-IR measurement. One way to improve the prediction accuracy is to use past data from the time series of an ingot growth. For the machine learning model, we used only the experimental values corresponding to the growth positions where the Oi concentration was measured, as shown in Fig. 1. However, the Oi concentration is also influenced by the history of the crystal growth process. Thus, it can be assumed that the time series data can improve the prediction accuracy, and this is currently undergoing study.

Figure 4 shows the measured and predicted Oi concentrations in a single ingot. The predicted values were calculated using actual values of the time series experimental data of the ingot. The prediction line indicates the fluctuation of Oi concentration along the ingot length. The result obtained demonstrates the high accuracy and usefulness of the machine learning model. Using the real-time prediction of the machine learning model, we established a system to predict Oi concentration instantaneously. Once the growth furnace was online, the process and monitored data were constantly input to the machine learning model. The data was then converted according to a specific format, so that it could be input to the machine learning model. The obtained prediction values are plotted in Fig. 4. The total time required in fetching an output was less than 1 s. Thus, we are able to know the Oi concentration at the position of the current growth interface immediately. In contrast, the conventional process of determining Oi concentration, starting from the crystal growth stage to the Oi inspection stage, spans across several days. This application will help reduce the time required to process the feedback of the material evaluation.
result that accelerates the development of materials with well-controlled target properties.

Unlike the discrete measured values, the predicted value is continuous along the horizontal axis denotingingot length. This is yet another advantage of using a machine learning model for prediction. Further, the continuous prediction of Oi concentration will lead to the detection of abnormal Oi fluctuation, which will prevent degradation of production yield due to the Oi concentration being beyond specification.

We made a machine learning model to predict Oi concentration in CZ silicon using the experimental data of approximately 450 ingots. The obtained prediction accuracy of RMSE $= 4.2 \times 10^{-6}$ atoms cm$^{-3}$ is high enough for practical use of the model for controlling Oi concentration. Using the real-time machine learning prediction model, we established a system that can instantaneously provide the Oi concentration at the position of the current growth interface. Further, we showed the importance of monitored and fixed parameters for accurate prediction and the advantages of parameter-influence analysis and continuous prediction using the machine learning model. These concepts will be useful for development of various other materials.

Acknowledgments This study was supported by the Japan Society for the Promotion of Science through KAKENHI Grant Number JP18K19033 and the Center for Advanced Intelligence Project, RIKEN. The authors acknowledge Kensaku Maeda of Tohoku university, Toru Ujihara of Nagoya university, Ichiro Takeuchi of RIKEN and Koji Izumome, Takashi Ishikawa, Hisashi Matsumura, and Hiroyuki Tsu atoja of GlobalWafers Japan Co., Ltd. for fruitful discussions.

References

1) S. Kishino, Y. Matsushita, M. Kanamori, and I. Izuka, Jpn. J. Appl. Phys. 21, 1 (1982).
2) K. Hoshikawa, H. Kohda, and K. Ikuta, Jpn. J. Appl. Phys. 20, 241 (1981).
3) S. M. Hu and W. J. Patrick, J. Appl. Phys. 46, 1869 (1975).
4) K. Sumino, H. Harada, and I. Yonenaga, Jpn. J. Appl. Phys. 19, L49 (1980).
5) T. Y. Tan, E. E. Gardner, and W. K. Rice, Appl. Phys. Lett. 30, 175 (1977).
6) K. Hoshikawa, H. Hirata, H. Nakanishi, and K. Ikuta, in Semiconductor Silicon, ed. H. R. Huff et al., (The Electrochemical Society, Pennington, NJ, 1981) , p. 101.
7) A. Murgai, in Semiconductor Silicon, ed. H. R. Huff et al., (The Electrochemical Society, Pennington, NJ, 1981) , p. 113.
8) T. Carlberg, T. B. King, and A. F. Witt, J. Electrochem. Soc. 129, 189 (1982).
9) H. Hirata and K. Hoshikawa, Jpn. J. Appl. Phys. 19, 1573 (1980).
10) R. E. Chaney and C. J. Varker, J. Cryst. Growth 33, 188 (1976).
11) K. Abe, K. Terashima, T. Matsumoto, S. Maeda, and H. Nakanishi, J. Cryst. Growth 186, 557 (1998).
12) X. Huang, K. Saitou, S. Sakai, K. Terashima, and K. Hoshikawa, Jpn. J. Appl. Phys. 37, 3188 (1998).
13) X. Huang, K. Saitou, S. Sakai, K. Terashima, and K. Hoshikawa, Jpn. J. Appl. Phys. 37, L193 (1998).
14) K. Kakimoto, K. W. Yi, and M. Eguchi, J. Cryst. Growth 163, 238 (1996).
15) K. Kakimoto, M. Eguchi, and H. Ozoe, J. Cryst. Growth 180, 442 (1997).
16) L. Liu and K. Kakimoto, Int. J. Heat Mass Transfer 48, 4492 (2005).
17) A. D. Smirnov and V. V. Kaliev, J. Cryst. Growth 310, 2970 (2008).
18) B. Gao and K. Kakimoto, J. Cryst. Growth 312, 2972 (2010).
19) B. Gao, S. Nakano, and K. Kakimoto, J. Electrochem. Soc. 157, H153 (2010).
20) R. A. Brown, T. A. Kinney, P. A. Sackinger, and D. E. Bornside, J. Cryst. Growth 97, 99 (1989).
21) T. K. Kinney, D. E. Bornside, W. Zhou, and R. A. Brown, in Semiconductor Silicon, ed. H. R. Huff et al., (The Electrochemical Society, Pennington, NJ, 1994) , p. 90.
22) D. E. Bornside, R. A. Brown, T. Fujiwara, H. Fujiwara, and T. Kubo, J. Electrochem. Soc. 142, 2790 (1995).
23) S. Takawgaya, Y. Shiraishi, K. Terashima, and S. Kimura, J. Electrochem. Soc. 142, 2839 (1995).
24) S. Takawgaya, Y. Shiraishi, K. Terashima, and S. Kimura, J. Electrochem. Soc. 142, 2844 (1995).
25) Y. Tsunooka, N. Kokubo, G. Hatasa, S. Harada, M. Tagawa, and T. Ujihara, Cryst. Eng. Commun. 20, 6546 (2018).
26) A. Boucetta, K. Kutsukake, T. Kejima, H. Kudo, T. Matsumoto, and N. Usami, Appl. Phys. Express 12, 125503 (2019).
27) Y. Takehara, A. Sekimoto, Y. Okano, T. Ujihara, and S. Dost, J. Cryst. Growth 532, 125437 (2020).
28) Y. Dang, L. Liu, and Z. Li, J. Cryst. Growth 532, 195 (2019).
29) K. Matsui, S. Kusakawa, K. Ando, K. Kutsukake, T. Ujihara, and I. Takeuchi, (2019), arXiv:1911.03671 [math-ph].
30) See [https://keras.io] (Accessed: 7 November 2020).
31) D. Kingma and J. Ba, (2014), arXiv:1412.6980 [cs.LG].

Fig. 4. (Color online) Measured and predicted Oi concentrations in a single ingot as function of ingot length. The measured data of this figure are included in the test data not in training data for the machine learning. In the real-time prediction system, the prediction line extends to the right as the ingot growth progresses, and we can obtain continuous profile of Oi concentration up to the current growth interface in the grown ingot.

K. Kutsukake et al.

© 2020 The Author(s). Published on behalf of The Japan Society of Applied Physics by IOP Publishing Ltd