Radiation Protection Dosimetry, 2023, **199(8–9)**, 767–774
https://doi.org/10.1093/rpd/ncad078
Paper

OXFORD

Deep TL: progress of a machine learning aided personal dose monitoring system

Evelin Derugin[1,*], Kevin Kröninger[1], Florian Mentzel[1], Olaf Nackenhorst[1], Jörg Walbersloh[2] and Jens Weingarten[1]

[1]Department of Physics, TU Dortmund University, 44227 Dortmund, Germany
[2]Personendosimetrie, Materialprüfungsamt Nordrhein-Westfalen, 44287 Dortmund, Germany

*Corresponding author: evelin.derugin@tu-dortmund.de

Abstract

Personal dosemeters using thermoluminescence detectors can provide information about the irradiation event beyond the pure dose estimation, which is valuable for improving radiation protection measures. In the presented study, the glow curves of the novel TL-DOS dosemeters developed by the Materialprüfungsamt NRW in cooperation with the TU Dortmund University are analysed using deep learning approaches to predict the irradiation date of a single-dose irradiation of 10 mGy within a monitoring interval of 41 d. In contrast of previous work, the glow curves are measured using the current routine read-out process by pre-heating the detectors before the read-out. The irradiation dates are predicted with an accuracy of 2–5 d by the deep learning algorithm. Furthermore, the importance of the input features is evaluated using Shapley values to increase the interpretability of the neural network.

Introduction

Handling sources of ionising radiation cannot be avoided in a variety of jobs in research, medical facilities and the industry. To ensure workplace safety in terms of adequate radiation protection measures according to regulations, people exposed to radiation above the background level due to their occupation may be required to wear a personal dose monitoring device or dosemeter[(1)].

In the event of an excess dose being registered on such a dosemeter, the on-site radiation protection officer oversees investigating the event or circumstance that led to the exposure to report on it and potentially enhance safety measures in place[(2)]. Despite digital dosemeters with real-time alarm functions being available, offering a certain degree of security enhancement, those come at additional cost and are not eligible for official dose monitoring in many countries[(3)]. As a result, workers may only carry a mandated passive dosemeter, often based on films (e.g.[(4)]), thermoluminescence (TL, e.g.[(5)]), optically stimulated luminescence (OSL, e.g.[(6)]) and other technologies, which do not allow for any real-time information.

In contrast to competing technologies, it has been shown for TL dosemeters that additional information about the circumstances of exposure can be derived at read-out time after a monitoring interval by analysing the detector signal, referred to as glow curve in Refs. ([(7–12)]). The best results reported by Mentzel *et al*.[(10)] achieve an estimation accuracy of the exposure date of up to 2 d within a monitoring interval of 1 month, using deep convolutional neural networks (CNNs). However, these studies of glow curves from the novel TL-DOS thermoluminescence dosemeters, developed by the Materialprüfungsamt NRW (MPA NRW) in cooperation with the Faculty of Physics of the TU-Dortmund[(13)], were performed under idealised laboratory conditions. The sensitive detector material is a thin layer of lithium fluoride doped with magnesium and titanium (LiF:Mg,Ti).

This work aims to extend previous studies in two ways with the goal to bring the technique closer to application in future routine dosimetry. First, an additional pre-heating step is included in the read-out process, which is required for routine dose monitoring using the TL-DOS system. This read-out process makes the dose estimation more robust (see, e.g.[(7)]), however, it reduces the time-dependent information within the glow curve signal significantly. To compensate for the loss of information, the number of measured glow curves is increased for training the algorithm.

Second, a new technique is introduced, focusing on deep learning-based algorithms' interpretability.

A common critique of machine learning predictions is the lack of insight into the reason for a particular prediction made by the model. In the case of radiation protection monitoring, the work of a responsible radiation protection officer can be greatly facilitated if the algorithm provides not only an estimate of the date of accidental exposure but also an interpretable output about the reliability of this prediction. This work takes a first step into that direction by using Shapley values to investigate how certain input features impact the prediction. First, a deep neural network (DNN) is trained on features manually derived from the glow curves and the impact of these variables on the prediction is investigated. This helps to understand which of the information obtained from the glow curves has a large impact on the prediction and gives insights into the irradiation scenario. Second, a CNN is trained directly on the 2D glow curve images to predict the exposure date. By analysing the importance of the image features, direct insight is obtained on which parts of the glow curves are crucial for the predictions and guide the visual inspections of the measured glow curves.

This paper is structured as follows: first, the novel data set is presented with a description of the glow curve shapes and the derived variables used for the DNN. Second, the prediction accuracy using the DNN is investigated and compared to previously published results before the Shapley value-based prediction interpretation method is introduced for this network. Afterwards, the 2D glow curve data and the corresponding CNN are introduced, their performance is evaluated, and the 2D result interpretation method is performed. Finally, the paper closes by discussing the presented approach's novelties and limitations and summarising the results.

Methods

Dataset for irradiation date reconstruction

The collected data of glow curves presented in the following is used for the first time in a publication. Compared to the measured data presented in a prior publication[10], this new measurement was performed using the current routine read-out process of TL-DOS and the latest read-out technology developed by the MPA NRW in Dortmund, Germany, for the future application in their dose monitoring service. The detectors were pre-heated at a temperature of $T_{\text{pre-heat}} = 431$ K for 9 seconds before the read-out process, as it will be done in the future routine dosimetry. Pre-heating enables a reproducible TL signal, which is required for the monitoring of occupationally exposed persons.

To record the data set, 4200 TL-DOS detectors were prepared for the irradiation campaign. The detectors are divided into 24 groups, with 150 detectors per group. Each group is prepared in such a way that, after the completion of the measurement campaign, each group experiences a different storage time between preparation and irradiation. The detectors were heated at $T_{\text{heat-out}} = 673$ K for 15 seconds to eliminate any residual signal and afterwards stored at constant room temperature of $T_0 = 292$ K. In the following 5 weeks, the groups were irradiated on different days with a single dose of $D = 10$ mGy using a Cs-137 irradiation source ($E = 662$ keV) and then stored again until the routine read-out process took place.

This procedure resulted in pre-irradiation storage times of $t_{\text{pre}} = [0, 1, 2, 3, 4, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 28, 29, 30, 31, 35, 36, 37, 38]$ d, representing the number of days between the preparation and irradiation of the detectors. After a total time of 41 d, corresponding to the time between preparation and read-out for 1-month monitoring interval including shipping times, the groups were read out in the same order as during preparation. The detectors were pre-heated at a temperature of $T_{\text{pre-heat}} = 431$ K for 14 seconds and then directly cooled by contact to a metal plate at room temperature ($T = 293$ K) for 3 seconds to interrupt the heating process. The pre-heated detectors were then read out at $T_{\text{read-out}} = 653$ K for 14 seconds.

Figure 1 shows the averaged glow curves for five different pre-irradiation storage times t_{pre}. The blue glow curve depicts an average curve for the detectors that were irradiated the same day after the preparation, whereas the purple curve represents the detectors that were irradiated 3 d before the read-out. Although it is rather a subtle effect, a change in the glow curve shape can be observed between the different times. Wich increasing pre-irradiation storage time, the first small peak nearly merges into the central one and the rightmost peak changes height over time. The apparent increase of the right peak results from the decrease of the central peak used to normalise the glow curve, as described in detail by Mentzel *et al.*[10].

Glow curve parameters

In a prior study[10], possible parameters that can be obtained from the glow curve in time–space were presented. These parameters are illustrated in Figure 2.

The region of interest (RoI) describes the range in which the actual thermoluminescence signal occurs. In routine dosimetry, the integral photon count within the RoI will be used for the dose estimation in the routine dosimetry. Other parameters can be derived from the RoI, such as the boundaries or its duration. The total and averaged number of all measured photon

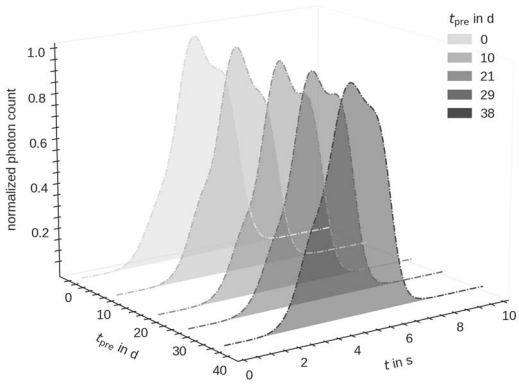

Figure 1. Glow curves for different pre-irradiation storage times. Each glow curve is an average of several glow curves and normalised by the respective maximum.

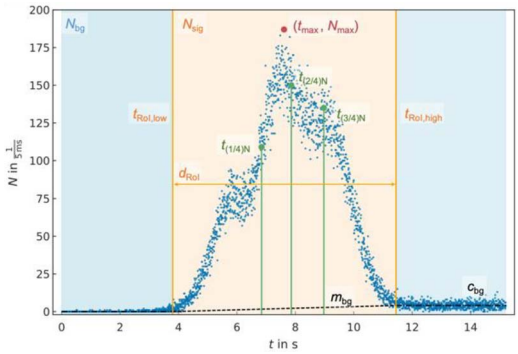

Figure 2. Illustration of a glow curve with the corresponding parameters in time–space. Reproduced with permission from Mentzel *et al.*[10].

counts is referred to as N_{tot} and N_{mean}. The maximum photon counts N_{max} is reached after t_{max} seconds. However, N_{max} can be neglected as a parameter for the training since the glow curves will be normalised to their respective maximum before they are transferred to the neural network to account for the inter-detector variation of the signal.

Furthermore, $t_{(i/4)}$ with $i = (1, 2, 3)$ marks the times at which the cumulative sum of photons within the RoI reaches the fraction $i/4$. The photon counts that are outside the RoI are considered as background and referred to as N_{bg}, where the average number of photons after the RoI is denoted as c_{bg}. The slope of the linear background within the RoI is denoted as m_{bg}. The consideration of the background parameters does not show any dependency towards the t_{pre} and are therefore removed from the input to the neural network. The total counts reduced by the background counts result in the signal counts N_{sig}. In total, 12 glow curve parameters are used as input features for training the neural network. All input variables are scaled to comparable ranges by shifting the mean to zero and scaling the distribution to unit variance. The target vector containing the corresponding values of t_{pre} is normalised to its maximum value.

The data are split into 70% training and validation data, which are used to adapt the weights and optimise the neural network's hyperparameter, and into 30% test data, which are used for an unbiased evaluation of the performance. The training and validation data set consists of $n = 2940$ individual glow curves. The hyperparameter optimisation is performed using a 5-fold cross-validation[19]. The fraction of samples of each detector group is kept equal in each of the sub-samples.

Machine learning models for irradiation date reconstruction

Both deep learning approaches are implemented using Keras[15] as the high-level Application Programming Interface (API) of Tensorflow 2[16]. Each model is optimised using a hyperparameter search based on the already-existing architectures from the prior study by Mentzel *et al.*[10]. To find the best set of hyperparameters, the mean-squared error (MSE) is calculated for each fold of the cross-validation and then averaged over all folds to take statistical limitations into account.

The final fully connected DNN, resulting from the performed hyperparameter search, consists of four hidden layers with a structure of [64, 128, 32, 16] nodes, each using the ReLU activation function[17], followed by a single-node output layer using a linear function. The network weights are adapted using the Adam optimiser[18] with a learning rate of $lr = 0.0007$ by minimising the MSE as a loss function and a batch size of 32. To prevent the model from overtraining, an L2-norm regularisation[20] with a strength of $\lambda = 0.0007$ is added to each hidden layer of the network.

Shapley values for feature interpretation

In addition to the performance of a machine learning model, a crucial element to ensure and especially understand the quality of the model is its explainability. A possible technique to allow for model interpretability are the SHapley Additive exPlanations (SHAP)[21], which aim to understand how input features contribute to the prediction using Shapley values[14].

Shapley values are normally used in cooperative game theory to fairly assign a player's contribution to the game's outcome. This principle can be transferred to a machine learning model by assuming that each input is a player in a game and the prediction is the outcome. SHAP enables the user to understand how a certain value of an input feature impacts the

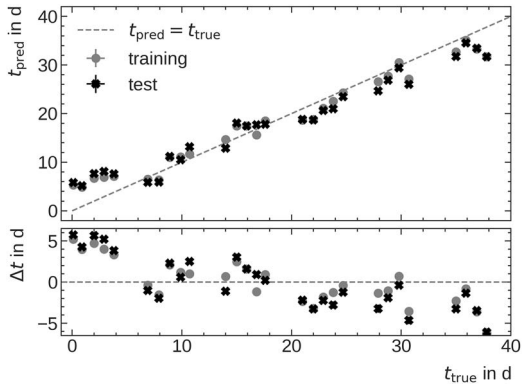

Figure 3. Averaged predicted t_{pre} for each measurement group with its standard error as a function of the true t_{pre} for the training (blue) and test (orange) data set is shown using the DNN model. The bottom plot shows the difference between the predicted and the true t_{pre}.

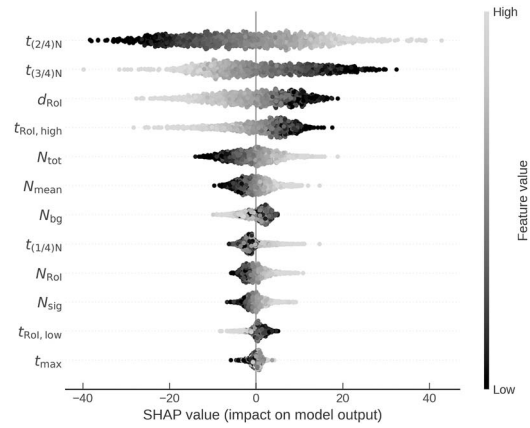

Figure 4. Local interpretation of the glow curve parameter with the corresponding contribution to the prediction of t_{pre}, calculated using the SHAP library.

prediction and provides an additive measure for feature importance.

Results

Irradiation date prediction with a DNN

After a successful hyperparameter optimisation and training of the final model architecture, the mean-predicted t_{pre} of each detector group and its standard error are shown in Figure 3 as a function of the true t_{pre} value as well as the relative deviation between the predicted and the true t_{pre} value for the training and test data sets. The average predictions for both training and test data do not deviate by more than 5 d from the true averaged t_{pre}. Due to the good agreement between the test and training data, a good generalisation capability of the DNN model can be confirmed. This means that the model does not memorise the training data but can interpolate well to unknown data.

Feature importance interpretation

The Shapley values for each glow curve parameter according to the predicted t_{pre} value are calculated using the *SHAP* library[21].

Figure 4 shows an exemplary composition of the contributions the individual glow curve parameters have on predicting the irradiation date for this detector. Starting from the global mean of all predictions, which is referred to as the base value (in this case, $E[f(x)] = 19.69$ d), the different parameters contribute individually to the final predicted value. Features coloured in red shift the prediction towards the positive direction (irradiation later within the monitoring interval) and features in blue towards the negative direction (earlier within the monitoring interval).

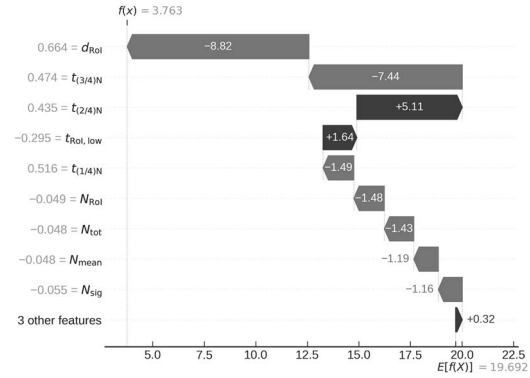

Figure 5. Summary SHAP plot with all glow curve parameters and their corresponding Shapley value in dependency of their feature value.

For example, the duration of the RoI has the most significant impact on the prediction in this case. The second largest impact on the prediction is the parameter $t_{(3/4)}$, whereas the other parameters only have a small impact on the prediction.

Apart from explaining for the individual predictions, the Shapley values can be combined with a global interpretation. Figure 5 shows all Shapley values aggregated from the prediction of all detectors within the data set. The features are sorted according to the Shapley values' maximum magnitude and their importance. The red–blue scale of the data points depicts the feature value from low to high.

As shown in Figure 4, the duration of the RoI has a high impact. Depending on the glow curve parameter, a tendency of the prediction can be derived. In case the duration of the RoI has low values, the prediction is influenced in a negative direction; in the case of

high values, the prediction shifts towards the positive direction. That means, on average, shorter glow curves indicate that the irradiation was received earlier within the monitoring interval. Other glow curve parameters show similar trends. These types of indicators might be already used by the radiation protection officer during visual inspection without analysing the glow curves in detail to gain insight into the irradiation event.

The most important features are the time stamps of $i/4$ of the cumulative sum of counted photons. As those are an indirect measure of the relative heights of the three visible peaks composing the glow curve, this indicates that the shape of the glow curve itself has high importance, which is investigated more closely in the next section.

Feature importance interpretation using a CNN

Pre-processing of the data and CNN architecture

As glow curves are usually displayed as 2D images to a user inspecting them, it is desirable to allow for a machine learning prediction explanation method that also operates on 2D image data in contrast to the previous parameter-based approach.

To do so, the glow curves are transferred into a 2D matrix and used as an image to train a CNN. This allows for Shapley value computation and assessment on a pixel-by-pixel base. A CNN uses trainable filter matrices, similar to common filter operations like edge detection known from traditional image processing, to extract features from the input data. In contrast to the common filter operations, CNN filter matrices are learned in the training process, allowing them to extract the most meaningful features for the prediction task that may not be obvious to a human user.

Several pre-processing steps are performed before transferring the glow curve as input to the CNN. First, the original measured glow curve is normalised to its respective maximum within an interval of [0, 100]. Second, an array with 100 entries is created for each measurement time. The corresponding index of the array that matches the normalised photon count number is set to the value 1. The remaining entries of the array are set to zero. This results in a 100 × 1320 pixel matrix representing the transformed glow curve, which is shown in Figure 6.

As this image was found to have unnecessarily many pixels not required for accurate predictions, the final pre-processing step is to reduce the dimension of the glow curve matrix by applying an average pooling layer. The best results were found using a window size of 4 × 20 pixels, resulting in a final 2D image size of 25 × 66 pixels.

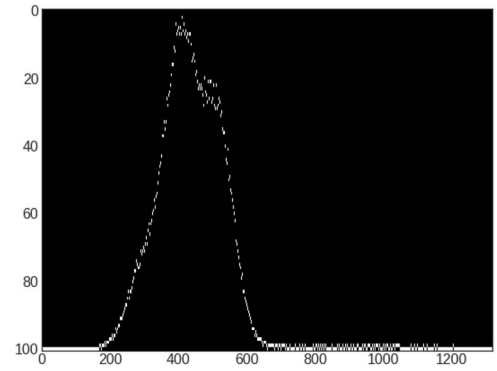

Figure 6. Representation of the transformed 2D glow curve.

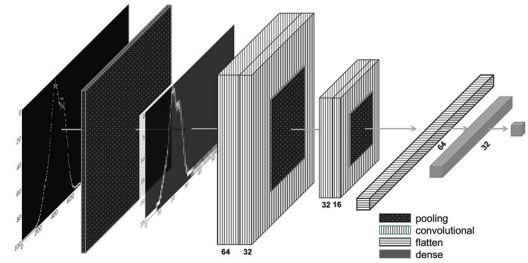

Figure 7. Schematic of the final CNN architecture with the pre-processing step.

A hyperparameter optimisation is performed using the same strategy as for the DNN to find the best model parameter of the CNN resulting in the best prediction results. The final CNN model architecture is depicted in Figure 7.

The model consists of two convolutional blocks each with two CNN layers responsible for the glow curve feature extraction followed by a dense-layer block combining the features for the pre-irradiation storage time prediction. The layers of the first convolutional block have 64 and 38 convolutional filters, and the layers of the second convolutional block have 32 and 16 convolutional filters. Both blocks have the same filter kernel size of (3 × 3) and the same activation function (ReLU). After each convolutional block, the network is regularised by a spatial drop-out function[21] with a drop-out rate of 0.15. After the second convolutional block, the output is flattened into a 1D array before passed to the dense block. The dense block consists of two dense layers with 64 and 32 nodes activated using a ReLU function and a single-node output layer activated using a linear function. To avoid overtraining and limit the effective capacity of the model an additional drop-out function[22] is applied after the first dense layer with a drop-out rate of 0.2. Furthermore, the model weights are optimised using the Adam optimiser again

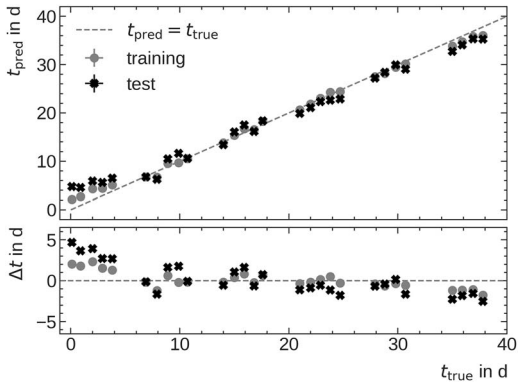

Figure 8. Averaged predicted t_{pre} for each measurement group with its standard error as a function of the true t_{pre} for the training (blue) and test (orange) data set is shown using the CNN model. The bottom plot shows the difference between the predicted and the true t_{pre}.

with an MSE error loss function, a learning rate $lr = 1e\text{-}4$ and a batch size of 16.

Irradiation date reconstruction and feature importance interpretation

Figure 8 depicts the predicted mean t_{pre} per detector group with the standard error as a function of the true pre-irradiation time t_{pre} for the test and training data using the final CNN model. The average predictions for training and test data are consistent mostly within 3 d compared to the true averaged t_{pre}. The prediction of the irradiation day with the CNN performs better than the DNN over the whole prediction interval, which is in agreement with the observations by Mentzel *et al.*[10]. The prediction of the irradiation day at very short and very long time spans is comparable to the rest of the monitoring interval. This is an improvement in the prediction accuracy compared to the DNN, which may be related to the manually derived glow curve input parameters. Because the CNN itself recognises and learns features within the glow curve, the prediction is no longer limited by the choice of input parameters. The uncertainties of the training, as well as the test data, agree well, suggesting that the model does not overtrain.

To understand the predictions obtained with the CNN better, the Shapley values are calculated using the DeepExplainer of the SHAP library[21]. As a result, all pixels of the 2D glow curve are provided with a Shapley value, which indicates the pixel's contribution to the prediction.

The Shapley values for two glow curves are exemplarily depicted in Figure 9. The red values increase the prediction value, while the blue ones decrease the prediction value. The most important pixels correspond

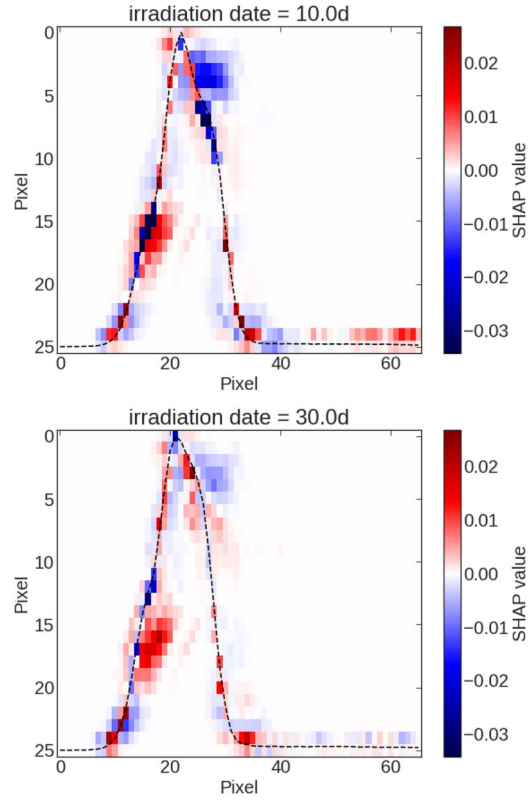

Figure 9. Two exemplary pre-processed glow curves at $t_{pre} = 10$ and 30 d with coloured pixels according to the individual pixel Shapley value. The black dashed line represents the glow curve progression.

to the features found to be the most important for the DNN prediction. For example, the pixels around (15, 15) and (5, 25) correspond to the values of the parameters $t_{(1/4)}$ and $t_{(3/4)}$, which are very relevant for the prediction using the DNN model. When comparing these two glow curves, a clear difference in the weighting of the pixels can be seen. For the curve with a t_{pre} of 10 d, pixels around the glow curve are coloured blue compared to the other curve, indicating that the prediction time is pushed to earlier times within the monitoring interval based on these pixel values. Furthermore, it can be confirmed that the glow curve pixels within the RoI have the most significant impact on the prediction, which confirms the impact of parameter d_{RoI}. Thus, the CNN is also interpretable in its predictions, which reinforces the use of this model in addition to the DNN.

Discussion

Both models, the DNN and the CNN model, achieve good prediction accuracy. While the DNN model yields

irradiation date predictions with uncertainties of up to 5 d, the CNN can predict the irradiation date mostly within 3 d. In contrast to the results obtained earlier by Mentzel *et al*.[10] in idealised lab conditions, these prediction accuracies are achieved using the routine read-out procedures, including the TL-DOS devices' pre-heating step reduces the differences in glow curve shapes for different irradiation dates significantly.

The models perform worst for irradiations in the first or last couple of days of the studied time interval, which indicates a bias towards the edges of the parameter space. However, for routine monitoring, these days are usually used for shipping the devices to and from the customer and are hence not part of the monitoring interval but can usually be neglected. However, the data used for these studies are still idealised in many aspects. Ongoing research hence focuses on expanding the presented techniques to more complex irradiation scenarios and more realistic storage conditions considering multiple irradiation doses or varying storage temperatures and other storage conditions, which are relevant in real-world dose monitoring scenarios.

Computing Shapley values and derived quantities for the machine learning models helps significantly to understand better how the models obtain their predictions. Furthermore, it enables experienced users with indicators for possible irradiation events upon visual inspection. By analysing the importance of features in the 2D glow curves, we could confirm the intuition about important shape features of the glow curves, which were formerly used as glow curve parameters. However, given that the CNN learns these features without the need for human knowledge and produces superior prediction accuracies because additional features are learnt, the glow-curve feature extraction seems unnecessary in future applications. In addition, it makes the model more flexible and robust for the prediction of additional information or changing irradiation or storage scenarios.

Conclusion

In this study, we present two deep-learning approaches for predicting the pre-irradiation storage time of pre-heated glow curves as they are recorded in the routine TL-DOS procedures. The best results are obtained with a 2D CNN model, which uses the glow curves as 2D images as input. This approach achieves on average prediction uncertainties of less than 3 d for the relevant monitoring interval.

Although the studies are performed using data recorded under laboratory conditions, such as the precise single-dose irradiation of the detectors with $D = 10$ mGy using a Cs-137 irradiation source and the storage of the detectors at a constant temperature over the whole monitoring interval, this work offers a first application of interpretable models on machine learning–aided radiation protection dosimetry. The offered interpretability allows for a user-sided assessment of the prediction and gives insights into the reasoning of the prediction. This aims at an easier use of the gained information by a radiation protection officer and a higher acceptance of the system due to a better understanding of the decision-making mechanisms, ultimately leading to improved radiation protection through successful personal dose monitoring.

Funding

This research was supported by the Deutsche Forschungsgemeinschaft (DFG), Project No. KR 4060/10-1.

References

1. Fantuzzi, E., Alves, J. G., Ambrosi, P., Janzekovic, H. and Vartiainen, E. *Implementation of standards for individual monitoring in Europe*. Radiat. Prot. Dosim. **112**(1), 3–44 (2004).
2. European Commission, Food And Agriculture Organization Of The United Nations, International Atomic Energy Agency, International Labour Organization, Oecd Nuclear Energy Agency, Pan American Health Organization, United Nations Environment Programme and World Health Organization. Radiation Protection and Safety of Radiation Sources: International Basic Safety Standards, IAEA Safety Standards Series No. GSR Part 3. (Vienna: IAEA) (2014).
3. International Atomic Energy Agency (IAEA). *Safety Standards Series No. GSG-7. Occupational Radiation Protection*. In: General Safety Guide. (Vienna: IAEA) (2018).
4. Mihai, F., Bercea, S., Stochioiu, A., Celarel, A., Udup, E. and Tudor, I. *Low and high dose measurement by Agfa personal monitoring film and FD-III-B badge dosimeter system*. Appl. Radiat. Isot. **68**(4–5), 614–616 (2010).
5. Kron, T., Lonski, P. and Yukihara, E. *Thermoluminescence dosimetry (Tld) in medicine: five 'W's and one how*. Radiat. Prot. Dosim. **192**(2), 139–151 (2020).
6. Yukihara, E. G. and McKeever, S. W. S. *Optically stimulated luminescence (OSL) dosimetry in medicine*. Phys. Med. Biol. **53**(20), R351–R379 (2008).
7. Theinert, R. Estimation of fading time and irradiation dose in thermoluminescence dosimetry using uni- and multivariate analysis techniques. Dissertation. (TU Dortmund, Germany) (2018).
8. Kröninger, K., Mentzel, F., Theinert, R. and Walbersloh, J. *A machine learning approach to glow curve analysis*. Radiat. Meas. **125**, 34–39 (2019).
9. Mentzel, F., Kröninger, K., Röhrig, L., Speicher, L., Steil, M.-L., Theinert, R. and Walbersloh, J. *Extending information relevant for personal dose monitoring obtained*

from glow curves of thermoluminescence dosimeters using artificial neural networks. Radiat. Meas. **136**, 106375 (2020).
10. Mentzel, F., Derugin, E., Jansen, H., Kröninger, K., Nackenhorst, O., Walbersloh, J. and Weingarten, J. *No more glowing in the dark: how deep learning improves exposure date estimation in thermoluminescence dosimetry*. J. Radiol. Prot. **41**, S506 (2022).
11. Walbersloh, J. and Busch, F. *A thin-layer Lif Thermoluminescence Dosemeter system with fast readout for the use in personal dosimetry services*. Radiat. Prot. Dosim. **170**(1–4), 191–194 (2016).
12. Isik, E., Isik, I. and Toktamis, H. *Analysis and estimation of fading time from thermoluminescence glow curve by using artificial neural network*. Radiation Effects and Defects in Solids **176**(9–10), 765–776 (2021). https://doi.org/10.1080/10420150.2021.1954000.
13. Weinstein, M., German, U., Dubinsky, S. and Alfassi, Z. B. *On the determination of the post-irradiation time from the glow curve of TLD-100*. Radiat. Prot. Dosim. **106**(2), 121–130 (2003).
14. Shapley, L. S. *A. value for n-person games*. Contributions to the Theory of Games **2**(28), 307–317 (1953).
15. Chollet, F. *et al*. Keras. GitHub. (available at: https://keras.io). (2015).
16. Abadi, M. *et al*. Tensorflow: large-scale machine learning on heterogeneous systems. arXiv:1603.04467. (available at: www.tensorflow.org/). (2015).
17. Nair, V. and Hinton, G. E. *Rectified Linear Units Improve Restricted Boltzmann Machines*. ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning. 807–814. (2010).
18. Kingma, D. P. and Ba, J. *Adam: a method for stochastic optimization*. Published as a conference paper at ICLR 2015. (2017) (arXiv:1412.6980v9).
19. Stone, M. *Cross-validatory choice and assessment of statistical predictions*. J. Royal Stat. Soc. **36**(2), 111–147 (1974).
20. Ng, A. Y. *Feature selection, l1 vs. l2 regularization, and rotational invariance Proc*. Twenty−First Int. Conf. Machine Learning, ICML '04, Association for Computing Machinery (New York, NY) pp. 78 (2004).
21. Lundberg, S. M. and Lee, S.-I. *A unified approach to interpreting model predictions*. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). (Red Hook, NY, USA: Curran Associates Inc.) 4768–4777. (2017).
22. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. *Dropout: a simple way to prevent neural networks from overfitting*. J. Mach. Learn. Res. **15**, 1929–1958 (2014).