scientific reports

OPEN

Atomistic and data-driven modeling of laser-induced graphene formation on sustainable polymer substrates

Cheol Hwan Kim[1], Jae Hyuk Kim[2✉], Sung-Yeob Jeong[3✉] & Bo Sung Shin[4✉]

Wood-based substrates—known for their renewability, abundance, and surface functionalization potential—have recently gained attention as polymers for laser-induced graphene (LIG) synthesis because of their environmentally friendly attributes. These environment-friendly properties also make them pollution-free and easy to dispose of after use. However, the formation of LIG on wood substrates lacks robust theoretical support, and molecular dynamics (MD) simulations, which are a potential theoretical framework, are time-consuming and computationally intensive. Herein, we employed temperature-dependent MD simulations to explore LIG formation on wood-based materials, validating our findings through a comparative analysis with atomic-scale characterization results. To address the high computational requirements of MD simulations, machine learning (ML) models, including long short-term memory (LSTM) networks, support vector regression (SVR), and multilayer perceptrons (MLP), were implemented to extrapolate predictions beyond direct simulation conditions. Each model exhibited high data explanatory power (R^2 values ≥ 0.9), and the computational time was significantly reduced compared to the MD simulations. ML-based predictions revealed a substantial correlation between the temperature and LIG formation extent, establishing an efficient framework for optimizing LIG synthesis from wood-based materials under various laser processing conditions. This framework has considerable potential for applications in energy storage devices, high-sensitivity sensors, and advanced catalytic materials.

Keywords Molecular dynamics, Machine learning, Laser-induced graphene, Wood

Laser-induced graphene (LIG) has recently gained prominence as a scalable and efficient technique for synthesizing graphene-like structures directly from carbon precursors[1–5], presenting a valuable alternative to conventional time- and energy-intensive methods such as chemical vapor deposition (CVD) and physical vapor deposition (PVD)[6–11]. The LIG process uses focused or unfocused laser beams to induce localized pyrolysis of carbon-based substrates, resulting in the formation of graphene-like carbon structures. This approach has been successfully demonstrated using various substrates, including polyimide (PI), biomass, and lignocellulose-based materials[12–16]. Polymer films containing PI have been extensively studied because of their high carbon content, excellent thermal stability, and ease of processing[17,18]. However, commercial polymer films have environmental and economic drawbacks owing to their nonbiodegradable nature, raising concerns about their sustainability and cost-effectiveness[19].

In response to these challenges, renewable lignin-based substrates, including wood, paper, and cork, have emerged as promising alternatives for LIG synthesis[20,21]. These materials offer advantages, such as low cost, availability, unique hierarchical structures, and the presence of functional groups that facilitate tunable porosity and surface modifications[13]. The unique composition of wood substrates, primarily cellulose, hemicellulose, and lignin, introduces distinct challenges and opportunities for the LIG process because these components undergo complex structural transformations under laser-induced pyrolysis. The formation of LIG typically occurs when carbon-based precursors are pyrolyzed at localized high temperatures (generally exceeding 2500 K)[22,23]. This high-temperature environment allows carbon atoms within the precursor to reorganize into a graphene-like

[1]Department of Cogno-Mechatronics, Pusan National University, Busan 46241, Republic of Korea. [2]Department of Industrial and Management Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. [3]Department of Mechanical Engineering, The University of Tokyo, Tokyo 113-8656, Japan. [4]Department of Optics and Mechatronics, Pusan National University, Busan 46241, Republic of Korea. ✉email: jaehyuk.kim@postech.ac.kr; ysjsykj8025@naver.com; bosung@pusan.ac.kr

structure, with the potential to further improve the layer quality and uniformity through the application of controlled atmosphere or inert gas environments during laser irradiation[14].

Despite significant advancements in the understanding of LIG formation mechanisms[24–28], there remains a lack of theoretical support for LIG synthesis from renewable materials. Molecular dynamics (MD) simulations using the reactive force field (ReaxFF) method offer a promising way to fill this gap[29], as they can represent the formation and destruction of interatomic chemical bonds. This provides detailed atomic-scale insight into these processes and allows for the tuning of laser parameters, such as temperature and pressure, to understand the mechanisms of short-time, microscopic reactions that are difficult to determine through conventional experiments. ReaxFF is particularly suited to capturing the complex processes of bond breaking and formation inherent in organic material pyrolysis as well as hybridization states and π-electron interactions essential for carbon-based systems[30,31]. These properties have enabled ReaxFF simulations to analyze the LIG formation process in carbon precursors effectively[23]. However, the high computational cost associated with ReaxFF simulations limits their practical applicability, particularly when exploring a wide range of processing parameters.

To address these limitations, this study integrated ReaxFF simulations with machine learning (ML) techniques to improve the computational efficiency and extend the predictive capabilities[32]. Temperature-dependent MD simulations can be applied to model LIG formation in the 2D carbon structures of wood substrates exposed to high-energy laser irradiation. The high computational load of the simulations can be overcome by training the simulation data with machine learning models and effectively predicting the carbonization process of wood substrates over time at different temperatures[33].

Three ML models—long short-term memory (LSTM), support vector regression (SVR)[34–36], and multilayer perceptron (MLP)[37–39]—were evaluated.

The results showed that the computation time for each model was up to 4 s, which was lesser than that of traditional MD simulations, and that the models were highly reliable, with an $R^2 > 0.95$, indicating that the results agreed well with the simulated values. This allowed us to identify the temperature range in which effective wood LIG could be achieved using pulsed laser processing. This framework, which combines MD simulations and ML, provides a comprehensive understanding of the underlying mechanisms driving LIG synthesis in renewable wood-based substrates and has the potential to optimize LIG applications under different laser processing conditions.

Materials and methods
Materials
Graphitized patterns on Meranti wood (Meranti wood, Indonesia) were created using a direct laser writing (DLW) technique. A 450 nm pulsed laser system (Laserpecker 2, Laserpecker, Shenzhen, China) was used with a scanning speed of 150 mm/s and repeated five times to generate the patterns. The specific laser specifications are presented in Table 1. Characterization of the patterns was conducted using Raman spectroscopy (RAMANtouch, Nanophoton, Tokyo, Japan). Atomic-scale analysis of the LIG structures was performed with a Cs-corrected scanning transmission electron microscope (Cs-STEM).

Molecular modeling
A molecular model that mimicked the structural characteristics of wood was developed by combining its three primary components: cellulose, hemicellulose, and lignin. A composite molecular model was constructed using the actual 5:2:3 mass ratio of cellulose, hemicellulose, and lignin to replicate the natural composition of hardwood[40]. Each molecular structure was geometrically optimized to minimize the energy and ensure structural stability before initiating molecular dynamics (MD) simulations.

Molecular dynamics simulation using ReaxFF
In this study, simulations were performed based on ReaxFF, using a parameter set applicable to C–H–O–N organic compounds used by Rahaman et al.[41] This parameter set can stably simulate the behavior of chemical reactions, bond formation and destruction, hydrogen bonding, dehydration reactions in various organic systems. Therefore the ReaxFF was used to model laser-induced pyrolysis and the subsequent formation of carbon ring structures in wood materials. The general energy equation is[42]

$$E_{system} = E_{bond} + E_{lp} + E_{val} + E_{over} + E_{under} + E_{coulomb} + E_{VdWaals}, \tag{1}$$

where E_{bond} denotes the bond energies, E_{lp} represents the lone pair energies, E_{val} represents the energies related to the valence angles, and E_{over} and E_{under} are the energies necessary to correct the over- and under-coordination

Laser type	NICHIA diode laser
Laser power	5 W
Laser wavelength	450 nm
Pulse frequency	2.5 kHz
Pulse width	5 ns
Laser spot size diameter	0.05 mm

Table 1. Detailed laser specification.

of the atoms, respectively. $E_{coulomb}$ represents the non-bonded Coulomb energy, and $E_{VdWaals}$ represents the van der Waals interactions.

The simulations were performed at a temperature range of 1000–4000 K within a simulation cell of dimensions 24.57 × 24.66 × 24.38 Å. The molecular dynamics simulations used a time step of 0.05 fs to fully capture the pyrolysis and carbonization processes, with each simulation running for 1 ns. During these simulations, the degradation of the cellulose, hemicellulose, and lignin matrices and the formation of carbon-based structures were monitored. The ReaxFF-based MD simulations were performed on the NURION system at the KISTI National Supercomputing Center.

Machine learning model for predicting graphene formation

In this study, LSTM, MLP, and SVR models were designed to predict outcomes based on time-series data generated from MD simulations. Since the task involves predicting the temporal evolution of molecular structures in response to temperature changes, a model that can intrinsically learn such time dependence is required. Although Convolutional Neural Network (CNN) could be considered as an alternative for time series prediction, CNN is primarily suited for spatial pattern recognition[43]. While LSTM, with its ability to model long-term dependencies, was selected as the primary model, its performance advantage over simpler models was found to be limited in practice.

To assess the necessity of temporal modeling, MLP and SVR—models that treat inputs as static feature sets—were included as computationally efficient baselines. These models offer fast training and simple implementation, making them suitable for comparison in scenarios where model interpretability and resource constraints are important. Compared to more complex algorithms such as XGBoost or random forests, which often require extensive tuning and longer training times, MLP and SVR provide a lightweight and practical alternative for initial benchmarking.

The primary variables were the results, temperature, and time step. The results are the surface areas of graphene obtained from the simulations, which are influenced by temperature and time.

The simulation data were analyzed independently under seven different temperature conditions with 2,001 time-series data points for each temperature. The model was validated using the k-fold method (k = 5) to ensure its reliability. k = 5 is a commonly used setting for small or medium-sized time series datasets, providing an appropriate balance between model learning stability and validation reliability. The data from the previous eight time steps for each temperature condition were used as the input characteristics, and the resulting value at the next time point was set as the prediction target. In addition, temperature data were integrated as auxiliary input variables into the LSTM model, which can be expressed as follows:

$$X_t = f(X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, X_{t-5}, X_{t-6}, X_{t-7}, X_{t-8}, U_t) \tag{2}$$

The entire dataset was divided into training, validation, and test sets in a ratio of 64:16:20. The input features are composed of graphene area units and were scaled to 0–1 using min-max normalization. Dimensionality reduction was not applied. The training data were used to train the model, the validation data were used for hyperparameter tuning and model optimization, and the test data were used to evaluate the generalization performance of the trained model.

Tables S1, S2, and S3 show the specific hyperparameter optimization processes for the LSTM, MLP, and SVR, respectively, obtained through the empirical experiments. The LSTM network consisted of three LSTM layers with 128, 128, and 128 neurons. The temperature input was combined with the output from the LSTM layers, and a linear activation function was employed to produce the final predictions. The performance of the model was assessed through 5-fold cross-validation with an epoch set to 50 by comparing the predicted and simulated results. A learning rate schedule of 0.05 was applied to fine-tune the optimization. The MLP used activation ReLU and set the random state to 512 and the maximum iteration to 500. SVR was used with Radial Basis Function (RBF) kernel.

Error analysis and evaluation

Three metrics, root mean square error (RMSE), mean square error (MSE), and Pearson correlation coefficient (PCC), which are widely used in regression model evaluations, were used for the evaluation.

RMSE is defined as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum\nolimits_{i=1}^{N} (y_{pi} - y_i)^2}, \tag{3}$$

where y_i is the target value for each sample set, y_{pi} is the predicted output, and N is the total number of sample points.

MSE is defined as follows:

$$MSE = \frac{1}{N} \sum\nolimits_{i=1}^{N} (y_{pi} - y_i)^2. \tag{4}$$

In addition, the mean absolute error (MAE) represents the average of the absolute values of the errors and is used to intuitively evaluate a regression model. It can be expressed as follows:

$$MAE = \frac{1}{N} \sum\nolimits_{i=1}^{N} |y_{pi} - y_i| \tag{5}$$

The performance of each model was evaluated using PCC[44], which measures the linear relationship between the predicted and simulated values. PCC values are between -1 and 1, with values closer to 1 indicating a stronger correlation between the prediction and simulated data. PCC is defined as

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \tag{6}$$

where r denotes the PCC, X_i represents the standard score for the target data, Y_i is the score for the output data, $\bar{X}$ is the targeted data mean, and $\bar{Y}$ is the output data mean. An overview of the entire process is presented in Fig. 1.

Results and discussion

Figure 2 shows the results of the molecular dynamics (MD) simulations of graphene formation at different temperatures (1000, 1500, 2000, 2500, 3000, 3500 and 4000 K). Video S1 shows the evolution of wood molecules over time at 3000 K. Each row corresponds to each temperature and shows the progression of carbon structure formation during laser-induced processing of a cellulose-hemicellulose-lignin model. In all temperature regions where sp^2 structures were formed, a transition from the initial molecule to amorphous carbon and back to graphene occurred for 1 ns. At a temperature of up to 1500 K, the carbon structure remained largely disordered through pyrolysis, and localized ring formation was observed, suggesting that there could be a temperature threshold for LIG formation using a pulsed laser. At 2000 K, regions with more organized ring structures began to form, indicating early stages of graphitization, whereas at 2500 K, larger and more distinct graphene regions were observed, with distinct and broadly varying ring planes throughout the structure. At 3000 K, well-defined amorphous graphene-like sheets were widespread, indicating that the graphitization process was significantly advanced and that extensive hexagonal carbon rings had formed. The progression up to 3000 K indicates that the temperature plays an important role in promoting the formation and stabilization of graphene within the molecular matrix, demonstrating a clear progression from disordered carbon structures to well-ordered graphene layers as the temperature increased. This indicates that as the temperature increases, the energy for
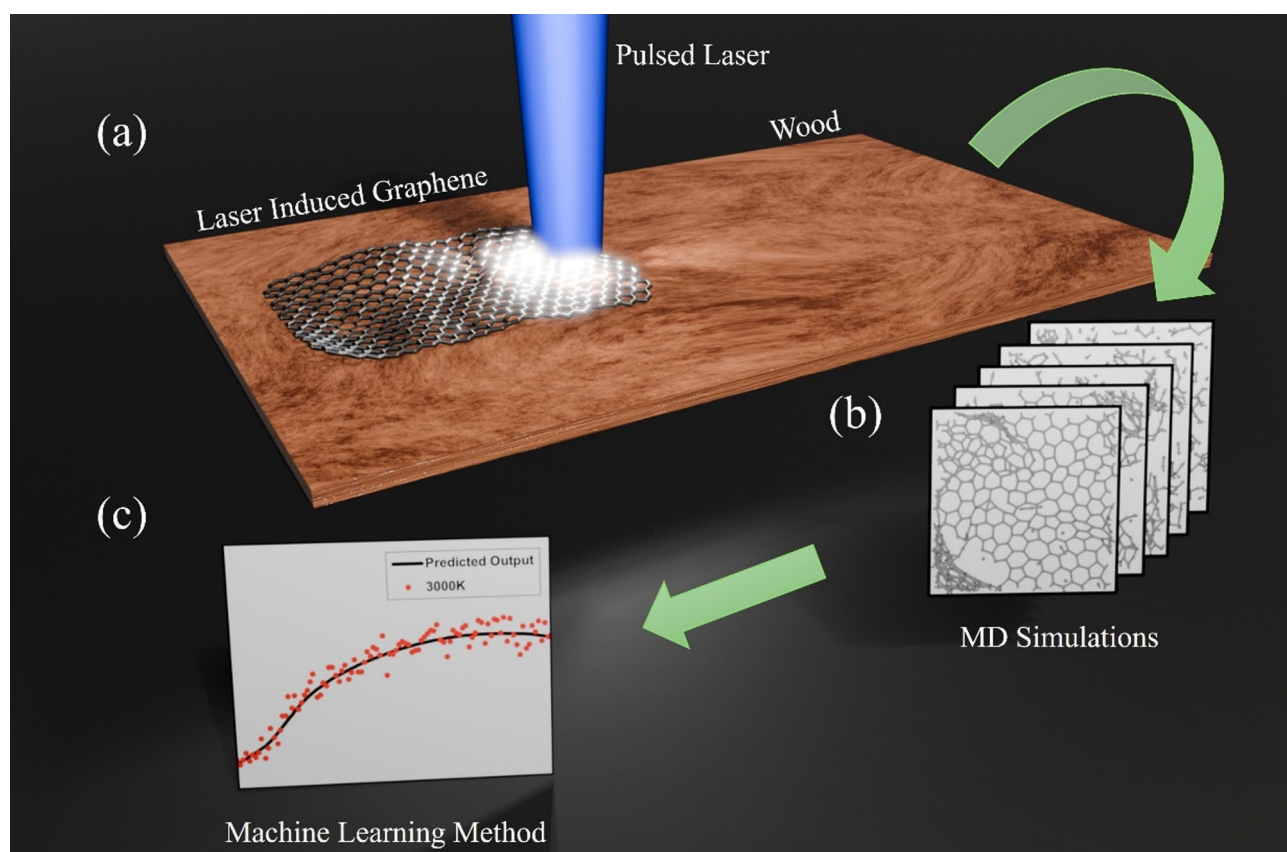

Fig. 1. Schematic of the fabrication process of wood-based LIG: (**a**) pulsed laser irradiates the wood surface, inducing the formation of graphene-like structures, (**b**) molecular dynamics (MD) simulations are performed to analyze the structural evolution, and (**c**) machine learning methods are employed to predict outcomes. The figure was created by the authors using Blender 4.2 (Blender Foundation, https://www.blender.org/).
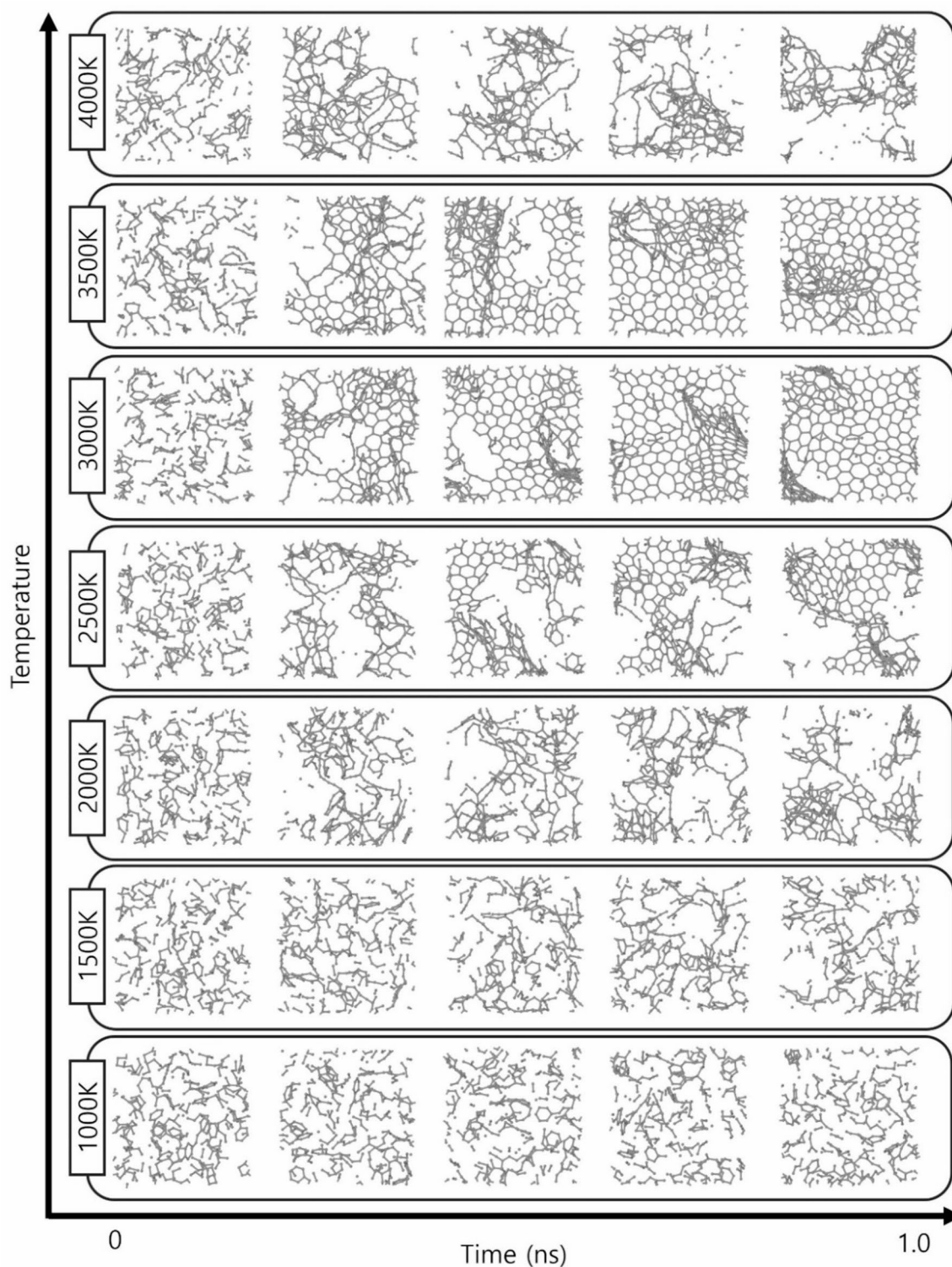
Fig. 2. ReaxFF simulation result showing the temperature-dependent evolution of graphene formation in a cellulose-lignin model under laser processing. The images represent the progression of carbon structure formation at temperatures of 1000, 1500, 2000, 2500, 3000, 3500, and 4000 K.

sp^2 hybridization increases. Since π-bonding occurred in this case, it could be expected that optimal electrical properties would be obtained at that temperature[45].

At 3500 K, the graphitization process continued, and a more localized sp^3 structure was observed than that at 3000 K. This indicates that temperatures above 3000 K exceed the optimum temperature for graphitization and that excess energy can lead to sp^3 hybridization. The process results in fewer π-bonds and more σ-bonds, which

destabilizes the two-dimensional carbon network and leads to the formation of an amorphous carbon structure. However, at 4000 K, the expected graphene structure hardly formed, and the carbon atoms aggregated to form a three-dimensional carbon structure. This phenomenon is consistent with the carbon pattern formation process described by Lee et al.[46]. These results show that high temperature conditions are not necessarily favorable for graphene structure formation, and based on molecular dynamics, we show that above 3500 K, the sp^2 carbon ring structure decreases and sp^3-based amorphous carbons predominate instead. These results provide new physical insights and suggest the importance of setting the proper fluence range for laser process optimization.

Figures 3a and S3 show the nanostructures of the isolated LIG flakes, with Cs-STEM images showing the polycrystalline structure of the wood LIG. The LIG flakes isolated from the substrate had a layered and corrugated morphology, with an abundance of five- and seven-membered rings as well as a six-membered ring structure of graphene. The rapid formation and cooling of this amorphous graphene using a pulsed laser resulted in the formation of a high-energy state structure. This process is referred to as kinetic graphene formation[12]. Figure S3b shows the characterization of the LIG. It shows several layers of structure with a d-space of 3.42 Å between graphene clusters. The morphology shown in Fig. 3b, which shows multiple carbon ring structures, is the result of the MD simulation of the sp^2 structure formation exposed to a temperature of 3000 K for 1 ns using a pulsed laser. Figure 3c is a magnified image of Fig. 3a obtained using Cs-STEM, showing the polycrystalline nature of the irregular LIG flakes. In particular, the five- and seven-membered rings observed in the simulation matches the defect structure in the TEM image, suggesting that ReaxFF qualitatively reproduces the microstructure formation and defect characteristics of carbon-based materials. This similarity provides a basis for ReaxFF to reliably simulate the graphitization process of wood-based precursors. These carbon rings are believed to be responsible for the D peaks observed in the Raman spectra. The presence of rings and defects enhances the electrical conductivity of graphene, supporting theoretical predictions for the manufacturability of LIG[47].

A large surface area is important for the effective application of LIG in sensors and catalysts. MD simulations were used to quantitatively analyze the structural evolution over time over a wide spectrum of temperatures, as shown in Figs. 4 and S1. Figure 4a shows the progression of the five-membered carbon rings, where the carbon atoms formed a pentagonal structure during LIG formation. With increasing temperature, the number of five-membered rings increased significantly, particularly at 3000 K. This increase suggests that certain temperatures enhanced the formation of carbon ring structures. However, the decrease in ring formation at 4000 K indicates that excessive heat altered the bonding structure of the structural organization, which is consistent with the results of studies on the thermal degradation of carbon-based materials under extreme conditions[23]. This indicates a temperature threshold for maintaining the stability of graphene formation and suggests that exceeding this limit can lead to defects and structural collapse. The sharp decrease at 4000 K indicates that such high temperatures can lead to the destruction of the sp^2 structures, suggesting that precise temperature control is required to ensure high-quality graphene formation. This is an important consideration when optimizing the laser parameters in LIG synthesis to maximize material performance. This variation in the rings with temperature was the same as that observed in Fig. 4b.
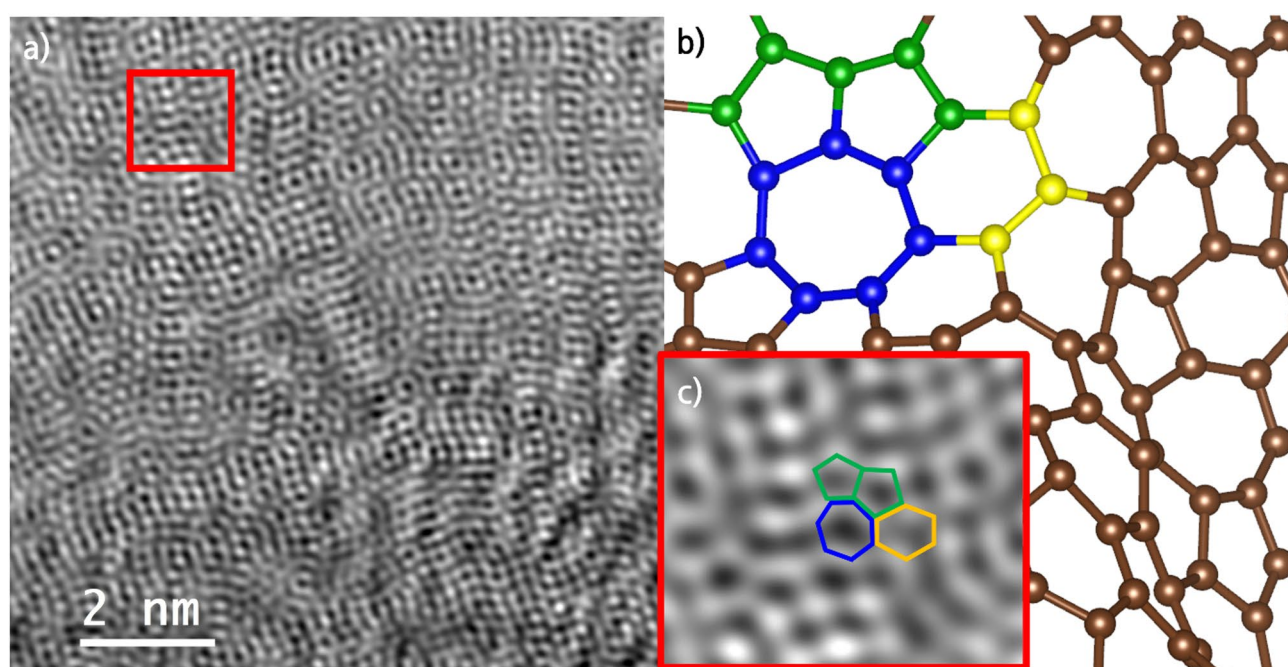

Fig. 3. (**a**) Cs-STEM image illustrating the polycrystalline structure of LIG. (**b**) Representations of the modeled configurations for five-, six-, and seven-membered carbon ring structures. (**c**) Enlarged view of the area outlined by the red rectangle in Fig. 3a.
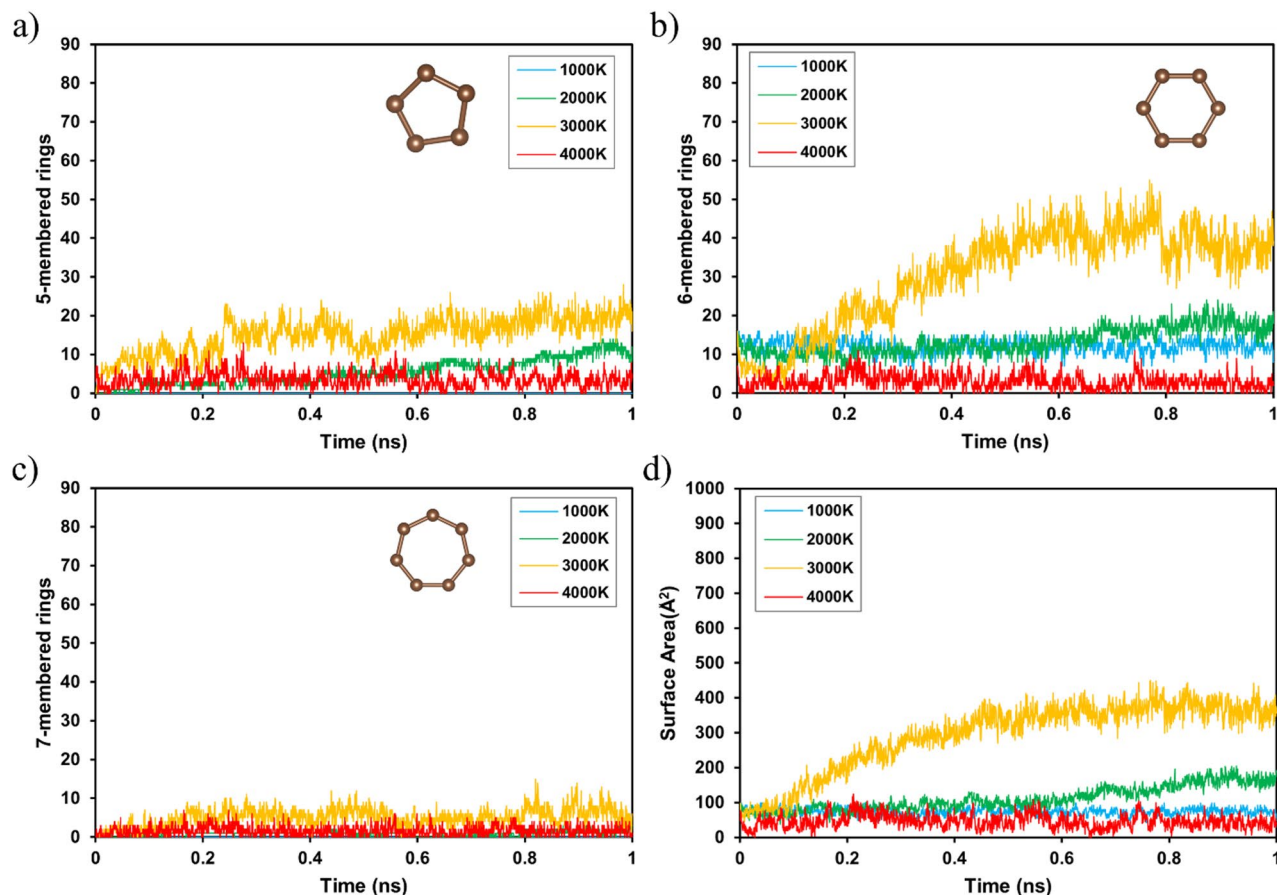
Fig. 4. Time-dependent changes of carbon ring structures and surface area during laser-induced graphene (LIG) formation in a cellulose-lignin model at various temperatures (1000, 2000, 3000, and 4000 K). (**a**) Evolution of five-, (**b**) six-, and (**c**) seven-membered carbon rings and (**d**) evolution of the surface area (Å^2).

In contrast, Fig. 4c shows the formation of seven-membered rings, which occurred less frequently than five- and six-membered rings. These rings are often associated with structural defects, as observed in graphene layers analyzed using Raman spectroscopy[47]. The results indicate that higher temperatures promote the formation of carbon networks; however, thermal energy can break the π-bonds that form between sp^2 carbon atoms. This leads to a decrease in the π-electron delocalization and reduces the flow of electrons, which can negatively affect the electrical conductivity of the material. LIG is often referred to as 3D graphene because it has a 3D graphitization structure owing to the torsion or shear stresses caused by the five- and seven-membered rings. Therefore, striking a balance between maximizing hexagonal ring formation and minimizing defects is a key challenge in optimizing the LIG process.

Figure 4d shows the change in surface area over time, which was calculated by adding the areas of the five-, six-, and seven-membered carbon rings at each time step. The increase in the surface area indicates that the carbon atoms were reorganized into larger and more extensive graphite layers. However, at 4000 K, the rate of surface area expansion decreased, implying that excess energy leads to aggregation rather than further development of organized graphene layers. The results indicate that temperatures around 3000 K promote graphitization; however, temperatures higher than this range may hinder the desired surface area expansion, thereby reducing the efficiency of graphene formation. These findings have important implications for the scalability of LIG synthesis, as they suggest the necessity to carefully optimize the processing temperature to balance material quality and structural integrity.

The gas molecules arising from the formation of LIG affect the porosity and electrical conductivity of graphene[30]. The molecules in the generated gas were counted based on the MD simulation results. Figures 5 and S2 show the temporal evolution of H_2 and H_2O molecules respectively during the formation of laser-induced graphene at different temperatures. Figure 5a shows the generation of H_2 molecules over time. It was observed that as the temperature increased, the number of H_2 molecules increased significantly, particularly at 3000 and 4000 K, with the highest generation at 3000 K. This suggests that at temperatures above 2000 K, the gas produced by promoting the degradation of the cellulose-hemicellulose-lignin matrix increased the formation of porous structures. This behavior is consistent with those observed in previous studies where high temperatures promoted the degradation of organic components, such as cellulose and lignin, which increased H_2 gas production during pyrolysis[48]. Due to the spatial limitations of MD simulations conducted on a nanometer
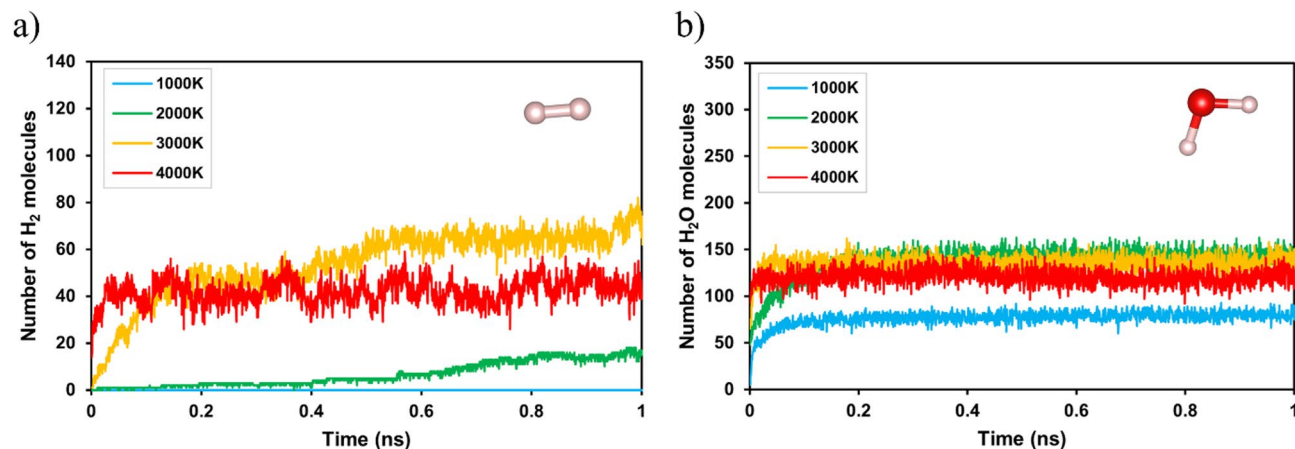
Fig. 5. Temporal evolution of the number of (**a**) H_2 and (**b**) H_2O molecules during laser processing of a cellulose-lignin model at various temperatures (1000, 2000, 3000, and 4000 K).

scale, there are limitations in quantitatively calculating the pore size distribution required for actual applications. However, gas molecules generated during the pyrolysis process could cause pore formation within the structure. Vashisth et al. confirmed that gas generation during LIG formation is correlated with the creation of porous structures[23]. When this was analyzed in conjunction with the structure shown in Fig. 2, the results show that a completely new structural characterization different from that of LIG emerged.

Figure 5b shows the production of H_2O molecules over time. In contrast to H_2, the number of H_2O molecules remained relatively stable over the temperature range, with a slight increase observed at 3000 and 4000 K. This suggests that temperature has less effect on H_2O production during LIG formation. Previous studies support this observation, noting that the formation of water during pyrolysis is less sensitive to temperature changes than H_2 gas production, which may be caused by the different chemical pathways of the decomposition of the biomass components[49].

To validate the effect of the graphene surface area on the wood material, the results were compared with the polyimide results at 3000 K. Because of the different cell sizes in the simulations, the volume differences were compensated for and calculated. Figure 6 shows a comparison of the time-dependent changes in the carbon ring structure and surface area of PI and wood at 3000 K during LIG formation. Figure 6a shows the evolution of the five-membered carbon ring over time. For the PI, the results were scaled to the same volume as that in the wood simulation. A gradual increase in the number of five-membered rings was observed for both PI and wood, with the PI generally exhibiting slightly higher values throughout the simulation. Figure 6b shows the evolution of six-membered carbon rings, which play an important role in the formation of graphene. Wood generally has a slightly lower number of six-membered rings compared to PI; however, in certain periods of the simulation, particularly after 0.5 ns, wood temporarily formed more rings, indicating that although PI tended to form more six-membered rings overall, wood exhibited stronger graphitization in certain segments of the process.

Figure 6c shows the evolution of seven-membered carbon rings. These rings were less prevalent in both models; however, PI exhibited slightly higher values, suggesting that the PI structure contained more defects. Figure 6d shows the evolution of the surface area over time. PI generally exhibits a slightly larger surface area than wood. The results show that PI tended to form a larger surface area overall and indicate the potential for trees to form LIG with fewer defects.

Figure 7 shows the Raman spectra as functions of laser fluence, with three peaks at 1337, 1573, and 2675 cm^{-1}, corresponding to the D, G, and 2D bands, respectively. These spectra indicate the formation of graphene-like carbon structures[50,51]. The D peak, which is related to sp^2 carbon bonds, indicated the presence of defects. The G peak represented the E_2g tangential vibrations of graphitic carbon, and the 2D peak provided evidence of a graphene-like flake structure, consistent with previous findings on LIG[52]. The experiments at all three fluences exhibited I_G/I_{2D} ratios greater than one, indicating the formation of graphene multilayers. The intensity ratio of the I_D/I_G band is a key indicator of the defect density and is affected by photothermal graphitization, which results in a graphene-like structure. As the laser fluence increased, the intensity of the G peak (I_G) increased, indicating enhanced graphitization under the laser treatment conditions[53].

At all fluences, I_D/I_G exhibited values less than 1, indicating that the LIG structure transitioned from amorphous carbon to a more crystalline graphene-like structure[45]. The increase in the G peak intensity, along with the 2D peak, suggests improved graphene quality, characterized by fewer defects and better layer alignment. This was supported by the decreases in defect density and I_D/I_G ratio, which indicated more complete photothermal graphitization. The results showed that the laser process successfully converted the precursor material into carbon, such as graphitized graphene, particularly at low scanning speeds or high laser fluences.

MD simulation results showed that graphene-like structures (sp^2 carbon ring structures) were clearly formed in the 2500–3500 K range, with the most developed form observed at 3000 K. Meanwhile, Raman analysis indicated that the G and 2D peaks increased with increasing fluence, indicating the formation of graphene structures with improved crystallinity. The fluence range at which these graphene characteristics began to be
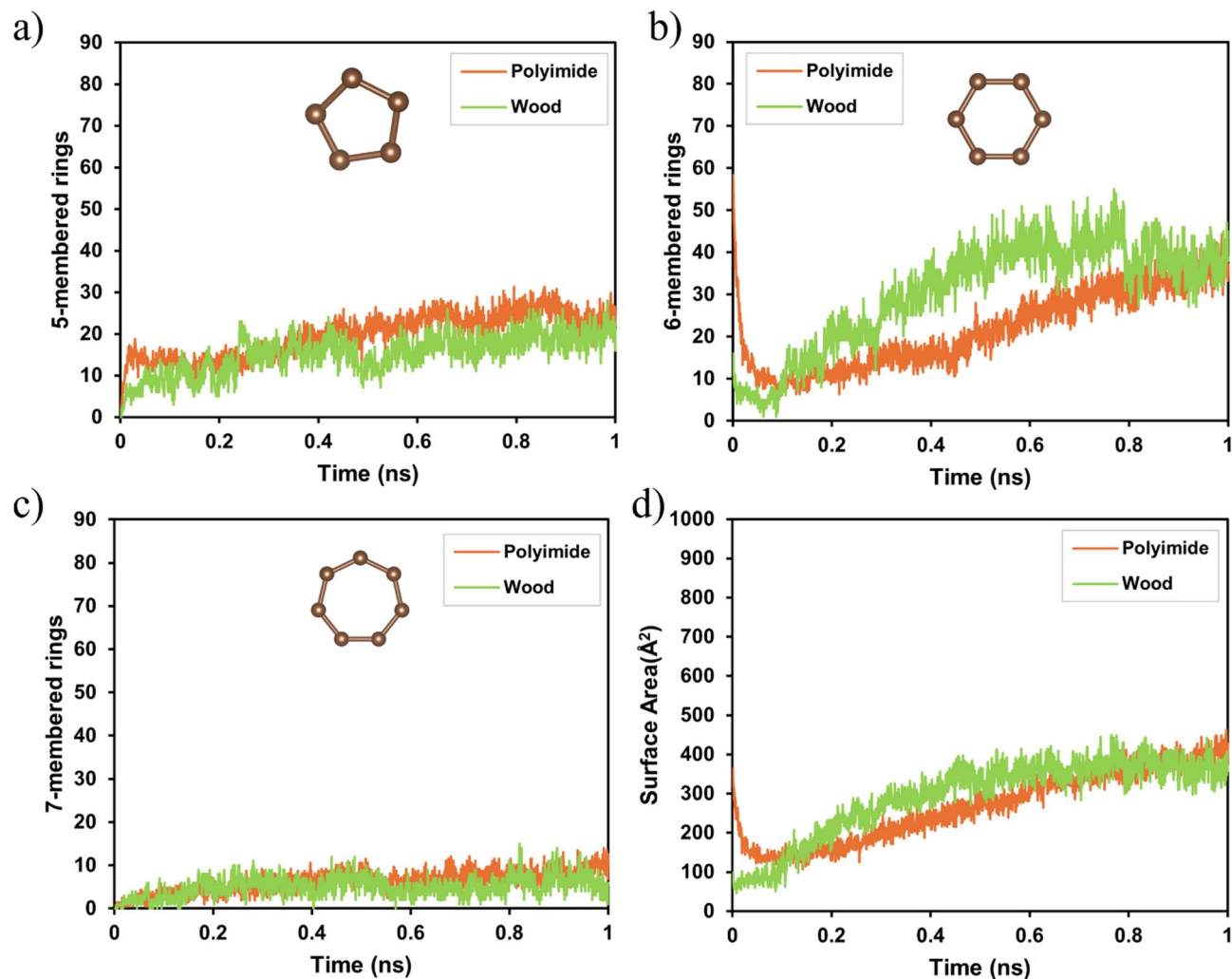
Fig. 6. Comparison of the temporal evolution of carbon ring formation and surface area expansion of polyimide and wood models during laser processing at 3000 K. The graphs show (**a**) five-, (**b**) six-, and (**c**) seven-membered rings, and (**d**) surface area ($\mathring{A}^2$) over time.

observed in the Raman spectra qualitatively matches the temperature conditions at which graphene structures began to form in simulations (2500 K or higher). This suggests that the results of the experiment and simulation are physically compatible.

The performances of the LSTM, MLP, and SVR models after machine learning were evaluated by quantitatively analyzing the models using metrics. Table S1-S3 summarizes the normalized performance metrics of the LSTM, MLP, and SVR models. Compared to the LSTM models, the MLP models had higher RMSE values and relatively lower R^2 values, suggesting that although the MLP models performed well, the LSTM models were better at capturing time series and complex nonlinear relationships because of their temporal nature. The SVR models had the highest RMSE and lowest R^2 values among the three models. However, the difference in the performance between the models was not significant. This limited performance difference is attributed to the relatively low complexity of the training dataset, which consists of a one-dimensional time series structure based on a single physical quantity. In addition, the simulation data shows a gradual change over time without abrupt transitions and tends to have little noise, resulting in similar performance across various models.

Figure 8 compares the performances of the three machine learning models—LSTM, MLP, and SVR—in predicting graphene formation during a laser-induced process. The scatter plot shows the relationship between the predicted output and simulated target value, with points closer to the red diagonal indicating more accurate predictions. The PCC, which measures the linear correlation between the predicted and target values, showed that all the models performed well. In general, the error tended to increase towards the end of time, which was attributed to the nature of the training data. In the latter part of the process, graphene was formed at high temperatures, and under periodic boundary conditions, graphene spanning the boundary of the cell was not counted as a ring and was represented as noise in the data. This caused a error in the scatter plots.

While all models were effective at predicting the data with PCC values greater than 0.97, the MD data used in this study was concentrated in specific temperature and time conditions, which limited the diversity of property
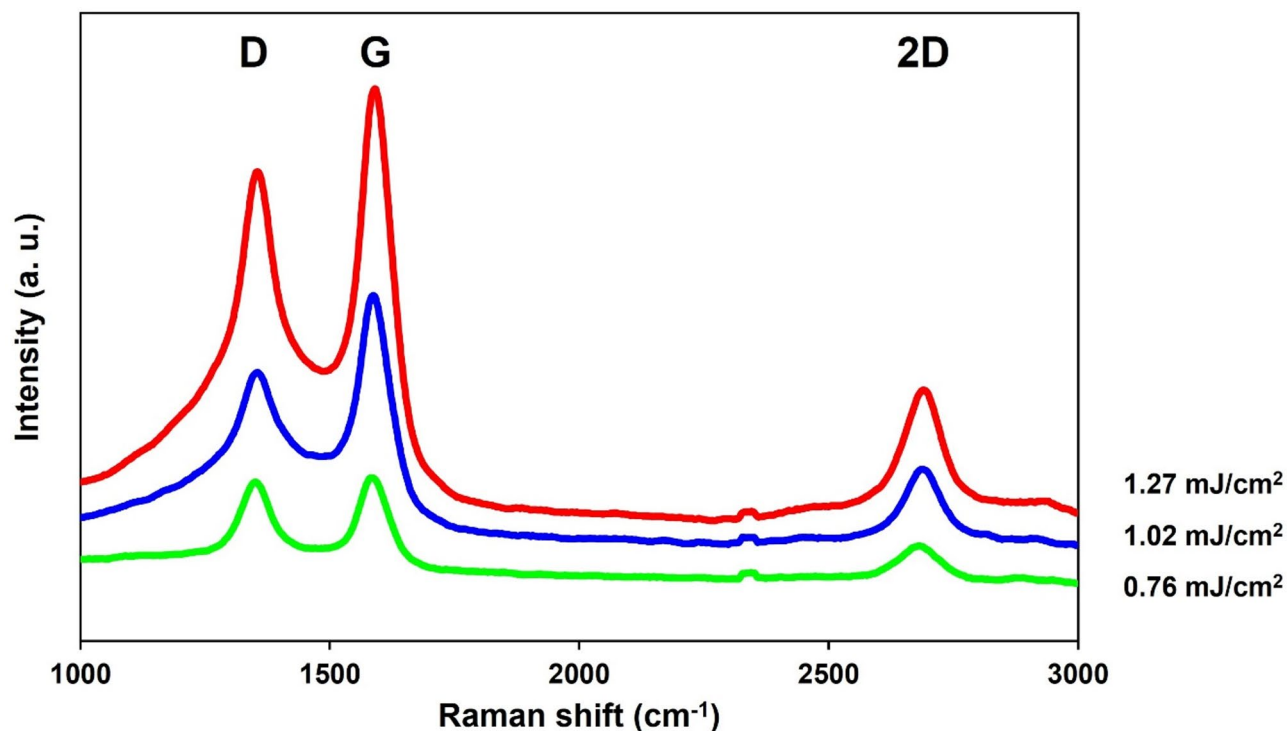
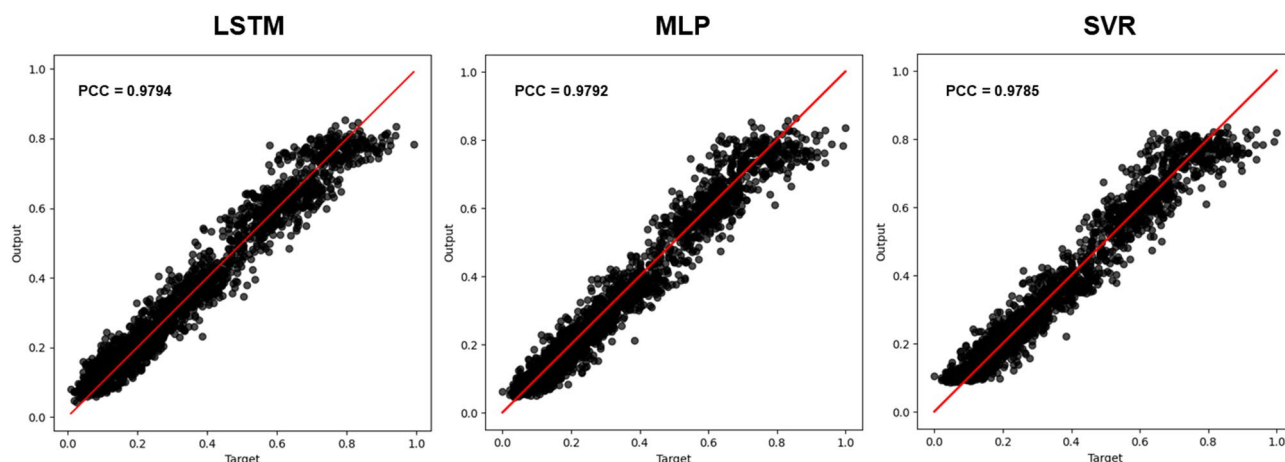**Fig. 7**. Raman spectra of the LIG irradiated laser at different laser fluences.

Fig. 8. PCC comparison of machine learning models (**a**) LSTM, (**b**) MLP, and (**c**) SVR in predicting graphene formation during laser processing.

values, making it difficult to distinguish performance differences between machine learning models. This could be explained by distribution bias, which tends to reduce prediction accuracy slightly in untrained conditions. Nevertheless, the LSTM model outperformed the other models with a PCC of 0.9794. The superior performance of the LSTM model is attributed to its ability to capture complex temporal patterns that are important in the laser-induced graphene formation process. Therefore, this predictive ability under limited data demonstrates the significant potential of the LSTM model. The MLP model exhibited similarly high PCC values. The SVR model, which is better suited to linear relationships, might have limitations in handling the nonlinear aspects of the data. Overall, these results show that machine learning is a useful tool for predicting graphene formation under different conditions and optimizing laser-induced processes.

These findings highlight the successful graphitization of the precursor material into graphene structures under various laser fluences. Based on this, further analysis was conducted to predict the evolution of the LIG surface characteristics using machine-learning techniques. Figure 9 shows the evolution of the validation and training losses over the course of training. The graph shows that the loss function decreased exponentially and
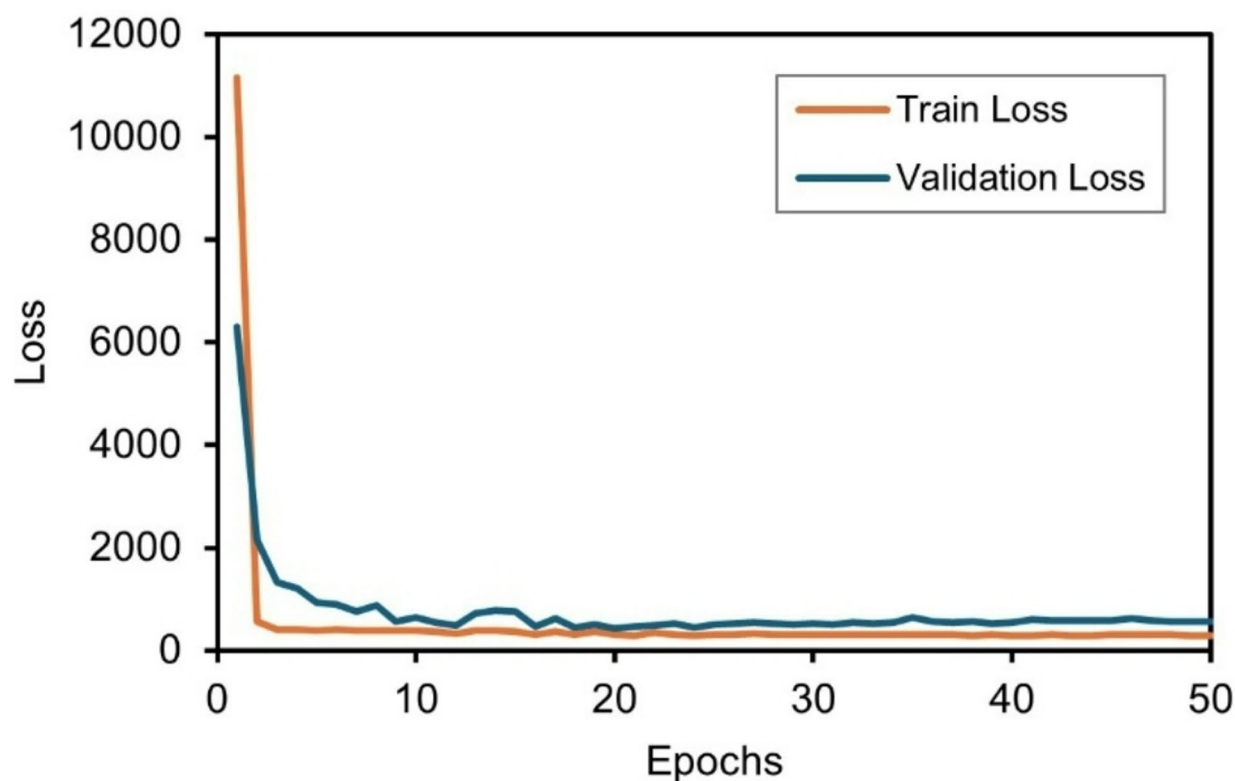
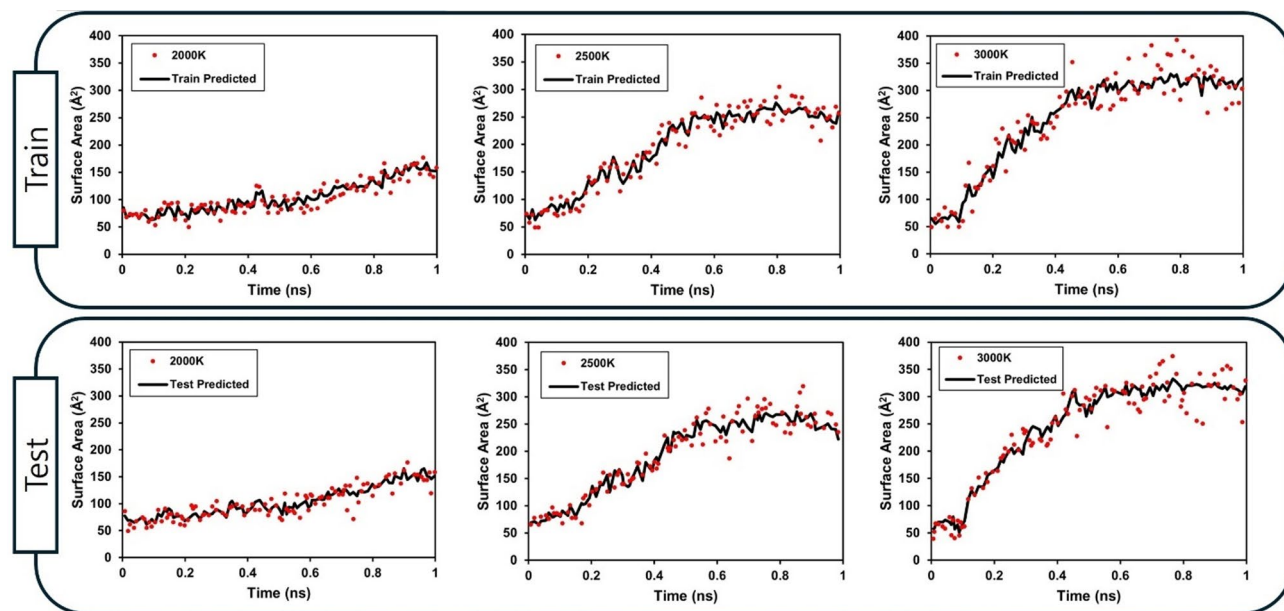**Fig. 9**. Loss in the training process in the LSTM model.

Fig. 10. Comparison of the predictions of the machine learning model LSTM for surface area formation over time on train and test data.

finally stabilized, indicating that it converged without overfitting or underfitting. Figure 10 shows the prediction performance of the LSTM model for the LIG surface area evolution at different temperatures by comparing the predictions for the training and test datasets and simulation results. The red dots represent the surface-area values obtained from molecular dynamics simulations, and the black lines denote the values predicted by the LSTM model. This model exhibited impressive generalization ability, particularly for curves where the true data

curve exhibits variations in each dataset. This indicates that the LSTM model accurately captures the formation of a structure over time.

The simulation was run in a configuration that used a total of 68 CPU cores in parallel, and a simulation corresponding to 1 ns (a total of 20,000,000 steps) was performed with a time step of 0.05 fs for a system containing 1,368 atoms. Under these conditions, the computational speed was 600,000 steps/day, resulting in a wall time of 800 h per temperature condition. This means that a computational resource of 68 cores × 800 h = 54,400 core-hours was required for a single simulation condition. On the other hand, machine learning-based predictions were performed in a general desktop environment (AMD Ryzen 5 3600), and the trained model required only about 4 s of computation time to perform one prediction. This has the effect of reducing computation time compared to MD simulations and has practical advantages in terms of being able to perform predictions quickly.

Conclusions

This paper presents a novel approach for understanding and optimizing the LIG formation process in wood-based materials by integrating MD simulations and ML techniques. A molecular model representing the natural components of wood (cellulose, hemicellulose, and lignin) was built to accurately simulate temperature-dependent pyrolysis and model the formation of 2D carbon structures. By precisely modeling the chemical reactions occurring during the decomposition of the wood matrix, ReaxFF was able to capture the emergence of carbon-based structures over the temperature range of 1000–4000 K. At certain temperatures, the formation of graphitic structures was accelerated, with clear evidence of increased formation of characteristic five-, six-, and seven-membered carbon rings.

The use of machine learning techniques, particularly the LSTM, SVR, and MLP models, significantly improved the efficiency of graphene formation prediction. Although the results were similar across the boards, the LSTM model exhibited the best performance, with the lowest RMSE and R^2 values compared to the MLP and SVR models. It also reduced the calculation time from tens to hundreds of hours in the MD simulation to a few seconds. This combination of MD simulations and machine learning is particularly noteworthy compared with traditional MD simulations, which are much more computationally expensive. By building a dataset that includes various variables, such as temperature, laser pulse cycle, pressure conditions, and material composition ratios, and expanding various reaction behaviors, such as gas generation, defect density, and hybridization ratio, as predictive variables to build a more high-dimensional and nonlinear dataset, it is expected that the characteristics between machine learning models will become clearer.

This integrated framework not only deepens our understanding of LIG formation mechanisms but also establishes a balance between MD simulations and machine learning models to ensure a comprehensive analysis of chemical reactions and process optimization. By combining the strengths of both methodologies, this approach provides a robust and scalable method for synthesizing graphene from renewable materials such as wood. The insights gained from this study can be applied to material selection and process parameter optimization to expand the applicability of the LIG technique. This study is expected to provide the foundation for the development of scalable and efficient graphene.

Data availability

The datasets used and/or analyzed in this study are available from the corresponding author upon reasonable request.

References

1. Carvalho, A. F. et al. Laser-induced graphene strain sensors produced by ultraviolet irradiation of polyimide. *Adv. Funct. Mater.* **28** (52), 1805271 (2018).
2. Jeong, S. Y. et al. Highly skin-conformal laser-induced graphene-based human motion monitoring sensor. *Nanomaterials* **11**(4), 951 (2021).
3. Sun, D. et al. Graphene-enhanced metal oxide gas sensors at room temperature: A review. *Beilstein J. Nanotechnol.* **9** (1), 2832–2844 (2018).
4. Chen, Y. et al. Interfacial laser-induced graphene enabling high-performance liquid–solid triboelectric nanogenerator. *Adv. Mater.* **33**(44), 2104290 (2021).
5. Le, T. S. et al. Green flexible graphene–inorganic-hybrid micro-supercapacitors made of fallen leaves enabled by ultrafast laser pulses. *Adv. Funct. Mater.* **32** (20), 2107768 (2022).
6. Wang, X. et al. Large-Scale synthesis of Few-Layered graphene using CVD. *Chem. Vap. Depos.* **15**, 53–56 (2009).
7. Munoz, R. & Gomez-Alexiandre, C. Review of CVD synthesis of graphene. *Chem. Vap. Depos.* **19**, 297–322 (2013).
8. Eigler, S. et al. Wet chemical synthesis of graphene. *Adv. Mater.* **25**, 3583–3587 (2013).
9. Cheng, Z. et al. Toward intrinsic graphene surfaces: A systematic study on thermal annealing and Wet-Chemical treatment of SiO2-Supported graphene devices. *Nano Lett.* **11**, 767–771 (2011).
10. Tahriri, M. et al. Graphene and its derivatives: opportunities and challenges in dentistry. *Mater. Sci. Engineering: C.* **102**, 171–185 (2019).
11. Zhang, X. et al. Mass production of biocompatible graphene using silk nanofibers. *ACS Appl. Mater. Interfaces.* **10** (27), 22924–22931 (2018).
12. Ye, R. & James, D. K. Tour. Laser-induced graphene: from discovery to translation. *Adv. Mater.* **31** (1), 1803621 (2019).
13. Le, T. S. et al. Ultrafast laser pulses enable one-step graphene patterning on woods and leaves for green electronics. *Adv. Funct. Mater.* **29** (33), 1902771 (2019).
14. Bressi, A. et al. Bioderived laser-induced graphene for sensors and supercapacitors. *ACS Appl. Mater. Interfaces.* **15** (30), 35788–35814 (2023).
15. Nam, H. et al. Smart wooden home enabled by Direct-Written Laser-Induced. *Graphene Adv. Mater. Technol.* **8** (9), 2201952 (2023).

16. Imbrogno, A. et al. Laser-induced graphene supercapacitors by direct laser writing of Cork natural substrates. *ACS Appl. Electron. Mater.* **4** (4), 1541–1551 (2022).
17. Ye, R. & James, D. K. Tour. Laser-induced graphene. *Acc. Chem. Res.* **51** (7), 1609–1620 (2018).
18. Bobinger, M. R. et al. Flexible and robust laser-induced graphene heaters photothermally scribed on bare polyimide substrates. *Carbon* **144**, 116–126 (2019).
19. Silvestre, S. L. et al. Cork derived laser-induced graphene for sustainable green electronics. *Flex. Print. Electron.* **7** (3), 035021 (2022).
20. Chyan, Y. et al. Laser-induced graphene by multiple lasing: toward electronics on cloth, paper, and food. *ACS Nano* **12**(3), 2176–2183 (2018).
21. Carvalho, A. F. et al. Laser-induced graphene piezoresistive sensors synthesized directly on Cork insoles for gait analysis. *Adv. Mater. Technol.* **5** (12), 2000630 (2020).
22. Jeong, S. Y. et al. Enhancing laser-induced graphene via integration of gold nanoparticles and titanium dioxide for sensing and robotics applications. ACS Appl. Materi. Interf. (2024)
23. Vashisth, A. et al. ReaxFF simulations of laser-induced graphene (LIG) formation for multifunctional polymer nanocomposites. *ACS Appl. Nano Mater.* **3** (2), 1881–1890 (2020).
24. Han, X. et al. Laser-induced graphene from wood impregnated with metal salts and use in electrocatalysis. *ACS Appl. Nano Mater.* **1** (9), 5053–5061 (2018).
25. Vaughan, E. et al. Sustainable Laser-Induced graphene electrochemical sensors from natural Cork for sensitive tyrosine detection. *Adv. Sens. Res.* **2** (12), 2300026 (2023).
26. Jung, Y. et al. Smart paper electronics by laser-induced graphene for biodegradable real-time food spoilage monitoring. *Appl. Mater. Today.* **29**, 101589 (2022).
27. Kulyk, B. et al. Conversion of paper and Xylan into laser-induced graphene for environmentally friendly sensors. *Diam. Relat. Mater.* **123**, 108855 (2022).
28. He, J. et al. Femtosecond laser One-Step direct writing electrodes with ag NPs-Graphite carbon composites for electrochemical sensing. *Adv. Mater. Technol.* **7** (11), 2200210 (2022).
29. Van Duin, Adri, C. T. et al. ReaxFF: a reactive force field for hydrocarbons. *J. Phys. Chem. A.* **105** (41), 9396–9409 (2001).
30. Ghavipanjeh, A. Simulation and experimental evaluation of laser-induced graphene on the cellulose and lignin substrates. *Sci. Rep.* **14** (1), 4475 (2024).
31. Kato, T. et al. Carbonization mechanisms of polyimide: methodology to analyze carbon materials with nitrogen, oxygen, pentagons, and heptagons. *Carbon* **178**, 58–80 (2021).
32. Jin, W. et al. Machine learning-based prediction of mechanical properties and performance of nickel–graphene nanocomposites using molecular dynamics simulation data. *ACS Appl. Nano Mater.* **6** (13), 12190–12199 (2023).
33. Noé, F. et al. Machine learning for molecular simulation. *Annu. Rev. Phys. Chem.* **71** (1), 361–390 (2020).
34. Smola, A. J., Bernhard & Schölkopf A tutorial on support vector regression. *Stat. Comput.* **14**, 199–222 (2004).
35. Li, M. et al. A deep learning convolutional neural network and multi-layer perceptron hybrid fusion model for predicting the mechanical properties of carbon fiber. *Mater. Design.* **227**, 111760 (2023).
36. Choi, C., Park, S. & Kim, J. Uniqueness of multilayer perceptron-based capacity prediction for contributing state-of-charge Estimation in a lithium primary battery. *Ain Shams Eng. J.* **14** (4), 101936 (2023).
37. Gardner, M. W. & Dorling, S. R. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmos. Environ.* **32**(14), 2627–2636 (1998).
38. Kanumuri, L. et al. Application of Support vector regression on mechanical properties of austenitic stainless steel 304 at elevated temperatures. *Mater. Today: Proc.* **2**(4), 1479–1486 (2015).
39. Long et al. Application of support vector regression to genome-assisted prediction of quantitative traits. *Theor. Appl. Genet.* **123**, 1065–1074 (2011).
40. Nhuchhen, D. R., Basu, P. & Acharya, B. A compressive review on biomass torrefecation. *Int. J. Renew. Erg Biofuel* **2014**, 1–57 (2014).
41. Rahaman, O. et al. Development of a ReaxFF reactive force field for Glycine and application to solvent effect and tautomerization. *J. Phys. Chem. B.* **115** (2), 249–261 (2011).
42. Chenoweth, K., Adri, C. T. & Van Duin Goddard. ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *J. Phys. Chem. A.* **112** (5), 1040–1053 (2008).
43. Li, Z. et al. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Trans. Neural Networks Learn. Syst.* **33** (12), 6999–7019 (2021).
44. Tang, M. et al. Performance prediction of 2D vertically stacked MoS2-WS2 heterostructures base on first-principles theory and pearson correlation coefficient. *Appl. Surf. Sci.* **596**, 153498 (2022).
45. Cooper, D. R. et al. Experimental review of graphene. *Int. Scholar Res Notices* **1**, 501686 (2012).
46. Lee, J. U. et al. Fabrication of UV laser-induced porous graphene patterns with nanospheres and their optical and electrical characteristics. *Materials* **13**, 3930 (2020).
47. Ferrari, A. C. & Denis, M. Basko. Raman spectroscopy as a versatile tool for studying the properties of graphene. *Nat. Nanotechnol.* **8** (4), 235–246 (2013).
48. Yang, H., Yan, R., Chen, H., Lee, D. H. & Zheng, C. Characteristics of hemicellulose, cellulose and lignin pyrolysis. *Fuel* **86** (12–13), 1781–1788 (2007).
49. Antal, M. J. & Gronli, M. The art, science, and technology of charcoal production. *Ind. Eng. Chem. Res.* **42** (8), 1619–1640 (2003).
50. Malard, L. M., Pimenta, M. A., Dresselhaus, G. & Dresselhaus, M. S. Raman spectroscopy in graphene. *Phys. Rep.* **473**, 51–87 (2009).
51. Ferrari, A. C. et al. Raman spectrum of graphene and graphene layers. *Phys. Rev. Lett.* **97**, 187401 (2006).
52. Eckmann et al. Probing the nature of defects in graphene by Raman spectroscopy. *Nano Lett.* **12**, 3925–3930 (2012).
53. Sun, Z. et al. Growth of graphene from solid carbon sources. *Nature* **468**, 549–552 (2010).

Author contributions

Funding

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-15945-2.

Correspondence and requests for materials should be addressed to J.H.K., S.-Y.J. or B.S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.