RESEARCH ARTICLE

LSTM Neural Networks With Attention Mechanisms for Accelerated Prediction of Charge Density at Onset Condition of DC Corona Discharge

YONG YI[1,2]**, (Senior Member, IEEE), ZHENGYING CHEN**[3]**, (Member, IEEE), AND RUI LI**[4]**, (Member, IEEE)**

[1]Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong, China
[2]Shenzhen Power Supply Company, Shenzhen 518020, China
[3]Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[4]Department of Electrical Engineering, Tsinghua University, Beijing 100083, China

Corresponding author: Rui Li (eeairicky@gmail.com)

ABSTRACT The onset process of corona discharge is naturally nonlinear and dynamic. The conventionally physical-based onset model and numerical computation of onset charge distribution are hampered by the computational power and given time. Here, in order to efficiently model this highly nonlinear dynamic process, a long short-term memory (LSTM) neural networks with attention mechanisms is proposed for accelerated charge density prediction under different atmospheric conditions, which adaptively choose charge-related input variables at each time step and hidden states relating to charge density all time steps. Our results demonstrate that this well trained model could make instant predictions with high accuracy under given target atmospheric conditions. Results show that the proposed model substantially reduces the computing time compared to physical-based methods. This work provides insights into applying LSTM neural networks to the charge density prediction of other discharge modes as well.

INDEX TERMS Neural networks, charge density, corona discharge, onset condition, prediction model, LSTM.

I. INTRODUCTION

Corona discharge is a nonequilibrium electrical discharge in the non-uniform electric field caused by the sharp electrodes. Corona discharge plays a vital role in a wide range of industrial applications, such as electrostatic precipitator [1], plasma reactor [2], ionic wind [3], plasma propulsion engine [4], or plasma spectroscopy [5]. The corona discharge commonly appears or forms when the surface electric field of the electrode reaches the onset electric field, which is commonly realized by the rod-plane electrode or the coaxial cylindrical electrode [6], [7], [8].

The associate editor coordinating the review of this manuscript and approving it for publication was Ilaria Boscolo Galazzo.

In recent years, to reduce or diminish the energy consumption of corona devices or suppress corona loss of high-voltage power apparatuses, the manipulation of corona onset voltage has attracted immense attention [9], [10], [11]. Predicting the onset characteristics and conditions through making use of the onset phenomenon becomes important for scientific research and industrial applications. In recent decades, there has been an extensive interest to investigate the onset electric field and physical-based onset models [12], [13], [14], [15]. These onset electric field formulas are approximately empirical expressions and derived from the experimental measurements, such as well-known Peek's formula [16]. However, these empirical formulas cannot capture the variation of the onset field with surrounding atmospheric condition
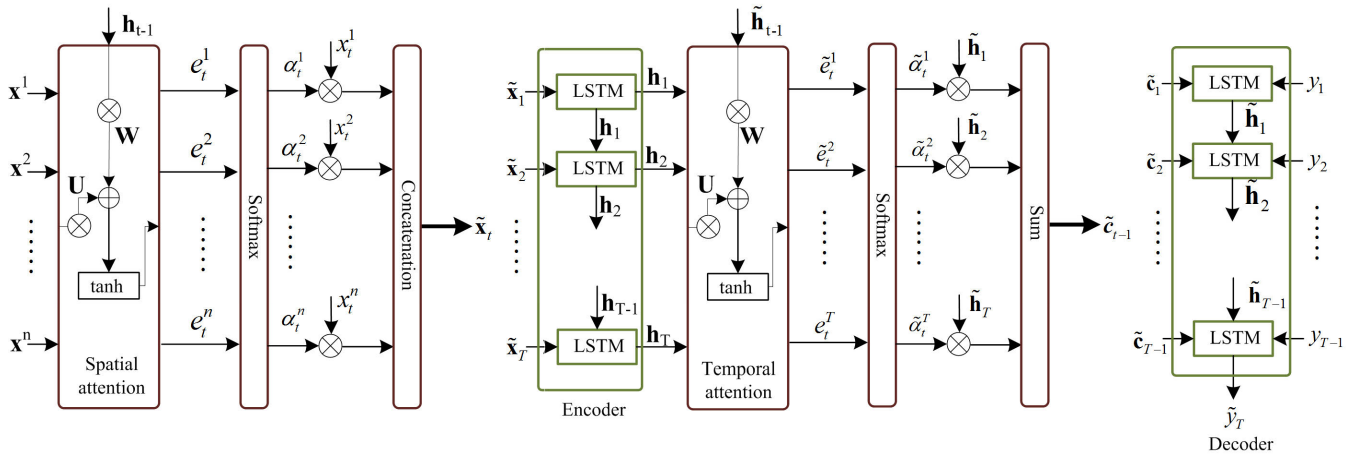
FIGURE 1. Illustration of the attention-based LSTM neural networks. The attention weight α_t^k is computed by the spatial attention (U and W are parameters), softmax layer and hidden state h_{t-1}. And then the new input $\tilde{x}_t$ is fed into the encoder LSTM cell. The attention weight $\tilde{\alpha}_t^k$ is computed by the temporal attention ($\tilde{U}$ and $\tilde{W}$ are parameters), decoder hidden state $\tilde{h}_{t-1}$ and softmax layer. The context vector $\tilde{c}_{t-1}$ is a sum of attention weights multiplied by decoder hidden state and fed to decoder LSTM cell. The $\tilde{y}_T$ is the prediction.

and electrode radius. The physical-based onset models are deduced from the avalanche growth, self-sustained mechanism and photoionization. Although the extensive computational efforts and some progress had been made, these methods were time-consuming and intense computing resources. The prediction of charge density properties was a highly nonlinear and complex problem. In addition, the atmospheric conditions (i.e., humidity, temperature, and pressure) have a great influence on corona onset. There are still many obstacles on the way to accurately predict the charge density during the onset process under different atmospheric conditions.

Machine learning, especially deep learning, has proven to be ubiquitous in a wide range of practical applications ranging from spoken words recognition to modern scientific research [17], [18]. Nowadays, researchers in the plasma or corona discipline are more willing to accumulate data with combinatorial experiments [19], [20] and physical model-based computations [21], [22]. Machine learning algorithms were deployed to find the potentially valuable pattern of plasma and predict the corona characteristics from these data [23], [24], [25], [26]. Some recent advancements applied neural networks to plasma discharge, such as plasma catalysis [27], [28], discharge process condition sensing [29], plasma chemical reactions [30], etc. Long short-term memory (LSTM) is a kind of recurrent neural network architecture, which is successfully applicable to nuclear plasma and partial discharge of apparatus. LSTM was used for controlling of plasma placement in Tokamak [31], density limit disruption prediction of Experimental Advanced Superconducting Tokamak [32], recognition of anomalous patterns of discharge in nuclear fusion [33], and plasma confinement mode classification [34]. In addition, LSTM had been applied in partial discharge detection and pattern recognition of insulated overhead conductors and gas insulated

switchgear [35], [36], [37]. Recent works have achieved significant improvements in predictions of discharge through neural networks. However, there are still few studies focusing on the prediction of onset characteristics under different atmospheric conditions and the application constraint of recurrent models to onset of corona discharge remains.

The main contributions of this article are as follows:

(1) Propose a long short-term memory (LSTM) neural networks with spatial and temporal attentions to model the charge density behaviors of DC corona discharge during the primary avalanche growth.

(2) Propose the Hausdorff metric to measure the similarity between the predicted distribution of charge density and the numerical calculation distribution of charge density.

(3) Analyze the spatial and temporal attentions and effectiveness of proposed model, which is compared with other traditional sequential models.

(4) Analyze the impacts of temperature, absolute humidity, and pressure on charge distribution of corona discharge, which aims at the reliable prediction of atmospheric impacts.

II. LSTM BASED AUTOENCODER ARCHITECTURE AND SPATIOTEMPORAL ATTENTION

A. MODEL ARCHITECTURE

The overall structure of the proposed method is the autoencoder architecture. The autoencoder architecture, including the encoder and decoder parts, is a competitive neural sequence transduction model and commonly proposed to address the sequential issue [38]. In the encoder framework, the encoder is built by stacked RNN, gated recurrent neural networks or LSTM. The input sequence is efficiently coded into the hidden state (dimensional vector) which encapsulating the useful information of input sequence. Then, the

decoder maps the encoding to the output sequence. The decoder is also built by RNN or RNN variants. The length of input sequence may differ from the output sequence.

For our study, Figure 1 presents the graphical illustration of the proposed attention-based LSTM neural networks for the sequence-to-sequence prediction, which belongs to the encoder-decoder architecture. To resolve the issue of different length between input and output sequences, the fixed-length context vector is produced from the hidden state and mapping function. In addition, when the length of input sequence increases, the performance of encoder-decoder neural networks deteriorates rapidly since it makes predictions based on the information of input and target sequences. To address this issue, the attention mechanism could be used for choosing the relevant context vector across all the time steps, which computes the representations of the input and output sequences.

As shown in Figure 1, in the encoder, the spatial attention is used to adaptively select the relevant input data related to charge density by referring to the hidden state at the previous time step. In the decoder, the temporal attention is used to automatically choose hidden states of the encoder related to the charge prediction across all time steps. Through these attentions, the LSTM neural networks could learn the long-term temporal dependency more effectively.

B. LSTM NEURAL NETWORKS

In the paper, the LSTM neural networks is used to construct the encoder-decoder framework. In each LSTM unit, a memory cell is used to determine what information could be passed or stored from input series by the input gate I_t, forget gate F_t, and output gate O_t. The long-term dependencies of the series are captured by the LSTM unit. The hidden state and update of the LSTM unit could be expressed as follows

$$F_t = \sigma\left(\mathbf{W}_{fx}\mathbf{x}_t + \mathbf{W}_{fh}\mathbf{h}_{t-1} + \mathbf{b}_f\right) \tag{1}$$

$$I_t = \sigma\left(\mathbf{W}_{ix}\mathbf{x}_t + \mathbf{W}_{ih}\mathbf{h}_{t-1} + \mathbf{b}_i\right) \tag{2}$$

$$O_t = \sigma\left(\mathbf{W}_{ox}\mathbf{x}_t + \mathbf{W}_{oh}\mathbf{h}_{t-1} + \mathbf{b}_o\right) \tag{3}$$

$$C_t = \tanh\left(\mathbf{W}_{cx}\mathbf{x}_t + \mathbf{W}_{ch}\mathbf{h}_{t-1} + \mathbf{b}_c\right) \tag{4}$$

$$\mathbf{h}_t = O_t \odot \tanh\left(F_t \odot M_{t-1} + I_t \odot (C_t)\right) \tag{5}$$

$$\mathbf{s}_t = F_t \odot \mathbf{s}_{t-1} + I_t \odot \tanh C_t \tag{6}$$

where C_t, $\mathbf{h}_t$ and $\mathbf{s}_t$ donate the intermediate state, hidden state and memory cell state of LSTM at time t, respectively; $\mathbf{x}_t \in \mathbb{R}^n$ represents the input series at time t and n is the number of input series; $\mathbf{W}_f$, $\mathbf{W}_i$, $\mathbf{W}_o$ and $\mathbf{W}_c$ represent the parameter matrices to be learned in LSTM; $\mathbf{b}_f$, $\mathbf{b}_i$, $\mathbf{b}_o$ and $\mathbf{b}_s$ represent the intercept parameters to be learned in LSTM; σ and tanh represent the sigmoid activation function and hyperbolic tangent function, respectively; $\odot$ represents the element wise multiplication.

C. SPATIAL AND TEMPORAL ATTENTIONS

The key challenge in charge prediction task (sequence transduction) is to learn the long-term dependencies. To improve the prediction performance for long sequence, the spatial and temporal attentions are used to learn the long-term dependencies.

For the given input time series $\mathbf{x}^k = \left(x_1^k, x_2^k, \cdots, x_T^k\right)$, the spatial attention mechanism could be constructed by the hidden state and memory cell state (at last time step $t-1$) of encoder LSTM. The spatial attention includes the multi-layer perceptron networks and softmax layer:

$$e_t^k = \mathbf{v}_e^k \tanh\left(\mathbf{W}_{es}\mathbf{s}_{t-1} + \mathbf{W}_{eh}\mathbf{h}_{t-1} + \mathbf{U}_e\mathbf{x}^k + \mathbf{b}_e^k\right) \tag{7}$$

$$\alpha_t^k = \exp\left(e_t^k\right) \bigg/ \sum_{i=1}^{n} \exp\left(e_t^i\right) \tag{8}$$

where e_t^k represents the output value of spatial attention, which measures the importance of input feature at time step t; $\mathbf{v}_e^k$, $\mathbf{W}_{es}$, $\mathbf{W}_{eh}$, $\mathbf{U}_e$ and $\mathbf{b}_e^k$ represent the parameters to be learned; α_t^k represents the output value of softmax layer.

$$\tilde{e}_t^d = \mathbf{v}_{\tilde{e}}^d \tanh\left(\mathbf{W}_{\tilde{e}\tilde{s}}\tilde{\mathbf{s}}_{t-1} + \mathbf{W}_{\tilde{e}h}\tilde{\mathbf{h}}_{t-1} + \mathbf{U}_{\tilde{e}}\mathbf{h}^d + \mathbf{b}_{\tilde{e}}^d\right) \tag{9}$$

$$\tilde{\alpha}_t^d = \exp\left(\tilde{e}_t^d\right) \bigg/ \sum_{m=1}^{T} \exp\left(\tilde{e}_t^m\right) \tag{10}$$

where $\tilde{e}_t^d$ represents the output value of temporal attention; $\mathbf{v}_{\tilde{e}}^d$, $\mathbf{W}_{\tilde{e}\tilde{s}}$, $\mathbf{W}_{\tilde{e}h}$, $\mathbf{U}_{\tilde{e}}$ and $\mathbf{b}_{\tilde{e}}^d$ represent the parameters to be learned; $\tilde{\alpha}_t^d$ represents the output value of softmax layer.

The final prediction output is

$$\tilde{y}_T = \mathbf{v}_y\left(\mathbf{W}_y\left[\tilde{\mathbf{h}}_T; \mathbf{c}_T\right] + \mathbf{b}\right) + b \tag{11}$$

where $\mathbf{c}_t = \sum_{d=1}^{T} \tilde{\alpha}_t^d \mathbf{h}_d$ is the context vector at time t; $\tilde{\mathbf{h}}_t$ is the hidden state of the decoder; $\left[\tilde{\mathbf{h}}_T; \mathbf{c}_T\right]$ is a concatenation of the context vector and the hidden state of the decoder; $\mathbf{v}_y$ and $\mathbf{W}_y$ represent the weight parameters; $\mathbf{b}$ and b are the bias vector and parameter.

D. EVALUATION MERICS

Three evaluation metrics (mean absolute error, root mean squared error and mean absolute percentage error) are used to measure the performance of onset characteristic prediction of different models. Specifically, the mean absolute error (MAE) is a scale-dependent metric and expressed as

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_t^i - \tilde{y}_t^i\right| \tag{12}$$

The root mean squared error (RMSE) is a scale-dependent metric and expressed as

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_t^i - \tilde{y}_t^i\right)^2} \tag{13}$$

The mean absolute percentage error (MAPE) is not a scale-dependent metric and given by

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_t^i - \tilde{y}_t^i}{y_t^i}\right| \times 100\% \tag{14}$$

where y_t^i and $\tilde{y}_t^i$ represent the label and predicted values at time t, respectively. N represents the total number of samples.

III. EXPERIMENTAL IMPLEMENTATION DETAILS

A. TRAINING AND TESTING DATASETS

The training and testing datasets are generated and collected from numerical computations and previous literature data [11], [12], [13], [14], [15], [16]. In our experiment, the number of avalanche charge is considered as the input variable. The 30000 labeled samples (number of primary avalanche and charge density) are collected during primary avalanche progression. The first 20000 samples are considered as training data. The following 2000 samples are used for validation data and the remaining 8000 are considered as the testing data. The input data is normalized to 0 to 1.

B. HARDWARE AND PARAMETER SETTINGS

The training and testing experiments are mainly implemented on a computing platform (NVIDIA RTX3090 GPUs with 24GB memory, Intel Core i9 10900K CPU @ 3.7 GHz and 32GB RAM), open-source PyTorch framework library, deep neural network library of NVIDIA compute unified device architecture (CUDA) and NVIDIA CUDA toolkit. In the comparison analysis experiments, the maximum 20 CPU threads are allocated for other non- neural network models. In addition, the NVIDIA compute unified device architecture (CUDA) toolkit and CUDA deep neural network (CUDNN) are used to accelerate the GPU performance. The PyTorch libraries, including numpy, matplotlib, sklearn, skimage, shutil, random, os, json, pandas, are used for the post hoc data processing and result visualization in our custom codes.

Firstly, in our LSTM model, there exist three main parameters to be optimized, including the numbers of hidden state in encoder and decoder parts and the number of time step of window. The grid search technique is used to determine the parameters which could achieve the best performance in the validation data. The number of time step of window is 10. The numbers of hidden states for encoder and decoder are all 64. In addition, the optimal minibatch size is 30 and the initial learning rate is set to 0.01. The number of epochs is determined as 60. In summary, our work employs the cross-validation procedure, learning-rate scheduler and automated early stopping criterion for training our proposed models to combat overfitting.

IV. PREDICTION RESULTS

Figure 2 (a) presents the prediction results and prediction residuals of charge density distribution with avalanche length of the rod-plane electrode under a specific atmospheric condition (the radius of the rod is 0.1 cm, temperature of 20 °C, pressure of 1 atm, relative humidity of 20 %; the scale type of the ordinate is linear). It is found that the predicted charge density distribution could fit the target curve very well. The statistical tests, such as error histogram and quantile–quantile plot, are used to validate the performance
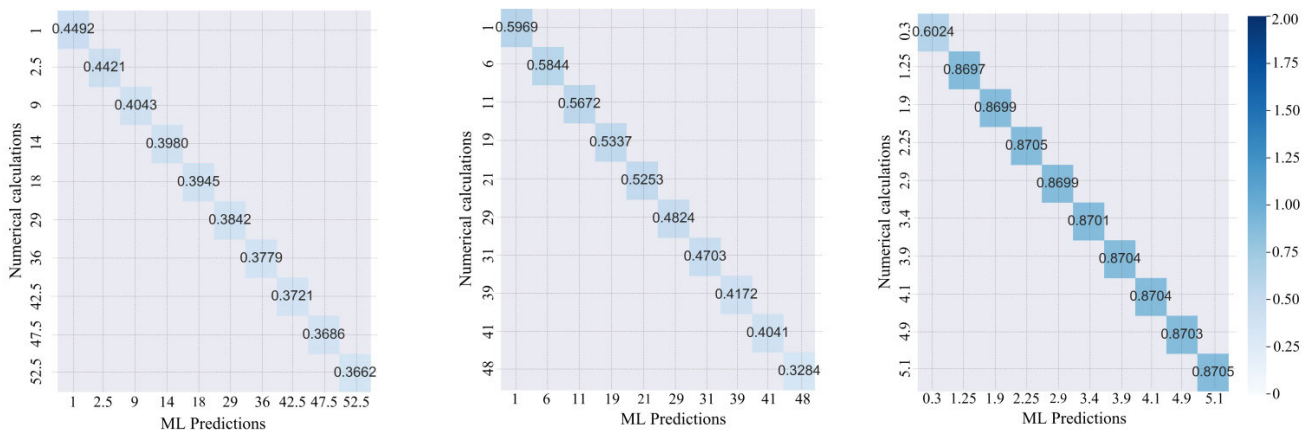

(a) Prediction of charge density and prediction residuals

(b) Histogram (c) Quantile-quantile plot
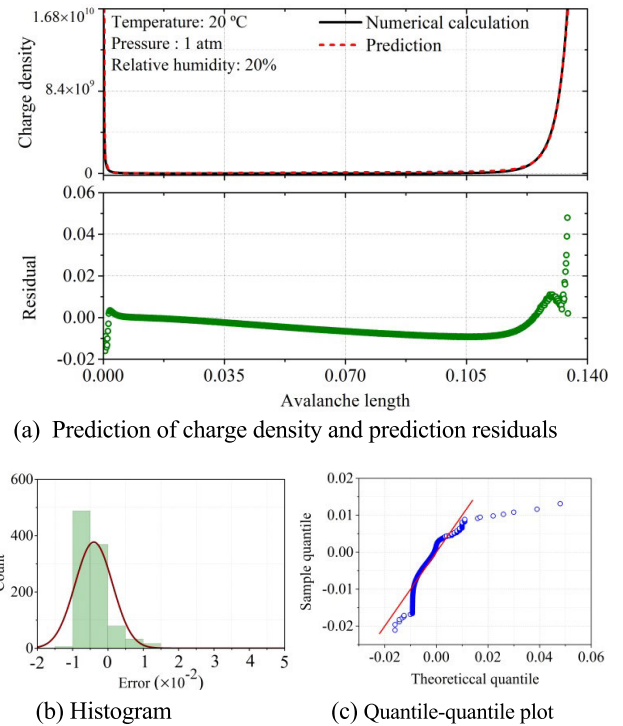
FIGURE 2. Model validation for the prediction of charge density (the radius of the rod is 0.1 cm, pressure of 1atm, temperature of 20 °C).

of proposed model. Figure 2 (b) and (c) present the error histogram and quantile–quantile plot of the prediction of charge density for the testing data. These statistical tests indicate that the prediction errors approximately obey Gaussian distribution. This quantile–quantile plot compares sample data on the vertical axis to normal distribution data on the horizontal axis, which suggests that the sample data are normal distribution.

To compare the prediction performance, the auto regressive integrated moving average (ARIMA), ARIMA model with external input (ARIMAX), recurrent neural networks (RNN), LSTM and the proposed model are used for prediction of charge density. The mean average error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE) are used as metrics of accuracy and these definitions are given in the supplementary material. TABLE 1 presents MAE, RMSE and MAPE of charge density prediction of negative glow corona discharge under temperature of 20 °C, pressure of 1 atm and relative humidity from 1 % to 52.5 %. We conduct the test 10 times for each model and present their average metrics and standard deviations in TABLE 1. It is observed that the performance of ARIMA and ARIMAX are lower than neural networks. Since ARIMA only considers the target series at previous time and ARIMAX is a linear model. The RNN could capture the nonlinear temporal relationships, and thus it outperforms ARIMAX. The LSTM neural networks could use a memory cell and activation function to learn long-term dependencies of the data for charge prediction. In addition, the additional comparative
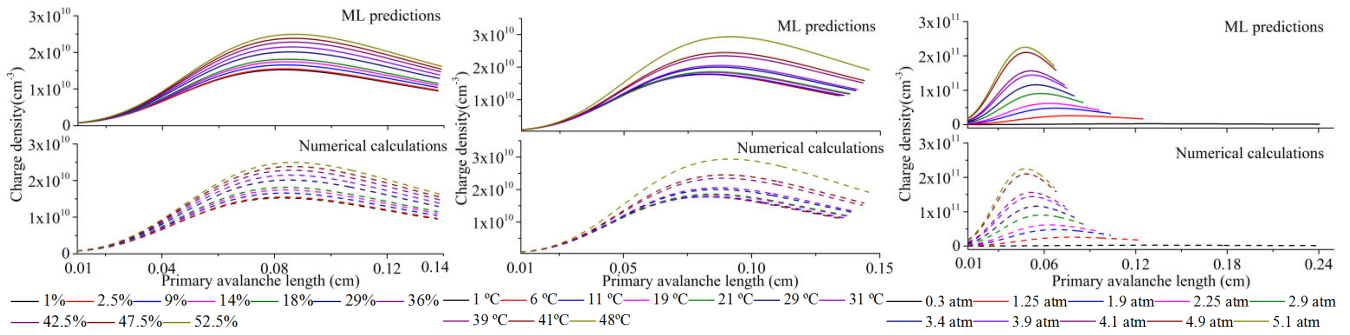
(a) Charge density under condition 1 (b) Charge density under condition 2 (c) Charge density under condition 3

(d) Hausdorff distance under condition 1 (e) Hausdorff distance under condition 2 (f) Hausdorff distance under condition 3

FIGURE 3. Comparisons of predicted and numerically calculated charge density at onset condition of negative glow corona discharge under different atmospheric conditions (the radius of the rod is 0.1 cm). Condition 1: temperature of 20 °C, pressure of 1 atm, relative humidity from 1 % to 52.5 %; Condition 2: pressure of 1 atm, relative humidity of 20%, temperature from 1 °C to 48 °C; Condition 3: temperature of 20 °C, relative humidity of 20 %, pressure from 0.3 atm to 5.1 atm. ML predictions represent the machine learning-based predictions.

analysis are performed on MAE, RMSE and MPAE of charge density prediction of negative glow corona discharge under the temperature from 1 °C to 48 °C and the pressure from 0.3 atm to 5.1 atm. We also conduct the comparisons on MAE, RMSE and MPAE of charge density prediction of positive glow corona discharge under the relative humidity from 7.5 % to 75 %, temperature from −25 °C to 65 °C and pressure from 0.25 atm to 5.75 atm. Our proposed model all significantly outperforms other four models. The results of our proposed model on the prediction of negative corona report the mean accuracies of 1.003 (MAE), 2.405 (RMSE) and 6.746 % (MAPE) under the pressure of 1 atm, relative humidity of 20%, temperature from 1 °C to 48 °C, and the mean accuracies of 1.454 (MAE), 2.625 (RMSE), 7.156 % (MAPE) under the temperature of 20 °C, relative humidity of 20 %, pressure from 0.3 atm to 5.1 atm. Our results of the proposed model on the prediction of positive corona report the accuracies of 0.464 (MAE), 2.333 (RMSE), 6.501% (MAPE) under temperature of 20 °C, pressure of 1 atm, relative humidity from 7.5 % to 75 %, the mean accuracies of 0.895(MAE), 2.709 (RMSE), 6.919 % (MAPE) under the pressure of 1 atm, relative humidity of 20%, temperature from −25 °C to 65 °C and the mean accuracies of 1.001 (MAE), 2.881 (RMSE),

7.182 % (MAPE) under the temperature of 20 °C, relative humidity of 20 %, pressure from 0.25 atm to 5.75 atm. The average computing overheads of ARIMA, ARIMAX, RNN, LSTM and well-trained proposed model are 25s, 33s, 22.5s 35s and 2s, respectively.

Our proposed model achieves the best performance among the five methods since it could extract the relevant information by spatial attention and then uses the temporal attention to choose the relevant information across all time steps.

TABLE 1. MAE, RMSE and MAPE of the five models for prediction result of charge density.

Model	MAE	RMSE	MAPE (%)
ARIMA	3.232±0.27	4.712±0.26	9.778±0.66
ARIMAX	3.013±0.23	4.011±0.22	9.101±0.36
RNN	1.891±0.19	2.997±0.17	8.899±0.24
LSTM	1.577±0.11	2.897±0.08	7.432±0.13
Our model	0.636±0.03	2.441±0.04	6.911±0.07
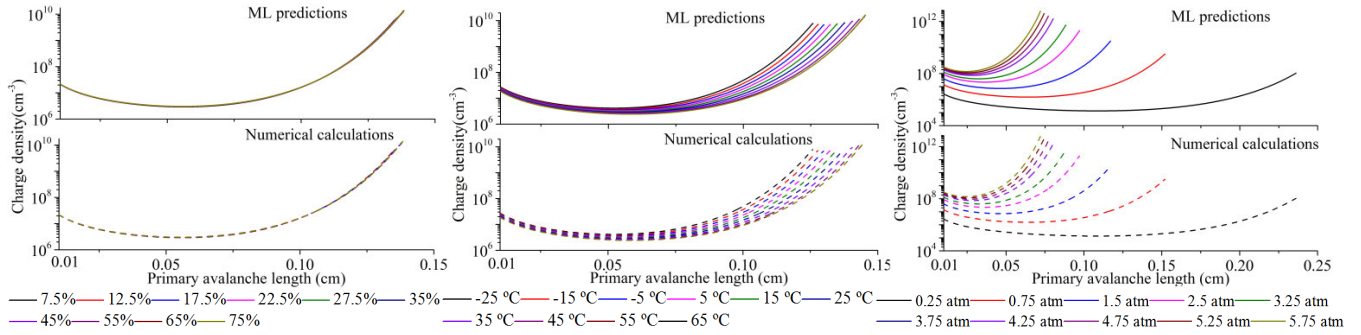
(a) Charge density under condition 4   (b) Charge density under condition 5   (c) Charge density under condition 6
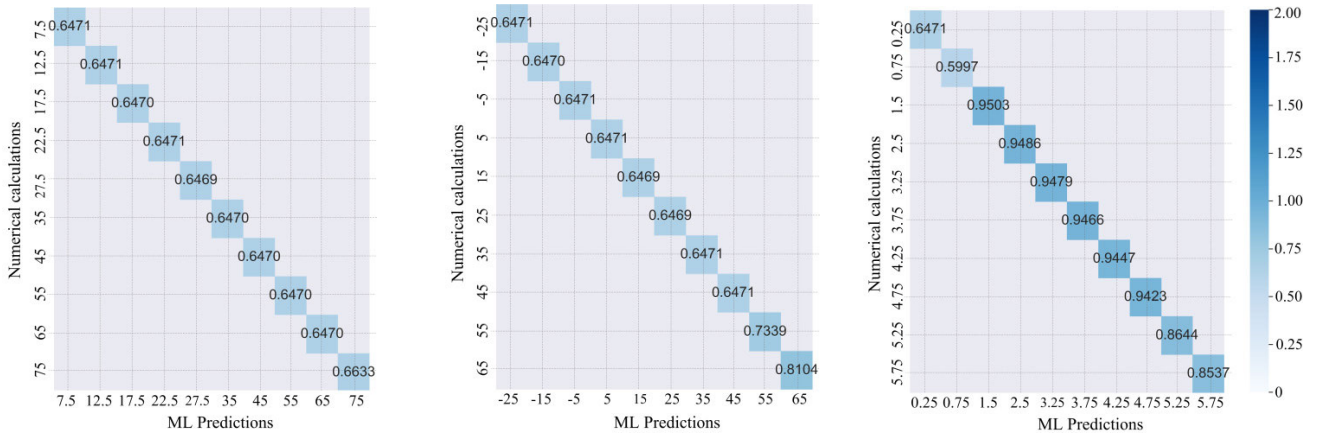
(d) Hausdorff distance under condition 4   (e) Hausdorff distance under condition 5   (f) Hausdorff distance under condition 6

FIGURE 4. Comparisons of predicted and numerically calculated charge density at onset condition of positive glow corona discharge under different atmospheric conditions (the radius of the rod is 0.1 cm). Condition 4: temperature of 20 °C, pressure of 1 atm, relative humidity from 7.5 % to 75 %; Condition 5: pressure of 1 atm, relative humidity of 20%, temperature from −25 °C to 65 °C; Condition 6: temperature of 20 °C, relative humidity of 20 %, pressure from 0.25 atm to 5.75 atm. ML predictions represent the machine learning-based predictions.

In Figure 3 (a)-(c), we present the predicted and numerically calculated charge density distribution with primary avalanche length of rod-plane electrode under different atmospheric conditions and the onset condition of negative glow corona. The solid lines represent the prediction results of the proposed model, and the dotted lines represent the numerical calculation results of physical-based onset model (the scale type of the ordinate is log10). For visual comparison, it is observed that the proposed model is in good match with the numerical calculations. As increasing the humidity, temperature and pressure, the predictions of charge density augment at the onset condition [24]. The predictive performance of our model is further test quantitatively by calculating curve similarity. Here the hausdorff distance (also called Hausdorff metric) is used to measure the similarity between two sets, which is a nonlinear operator and quantifies how the predicted curve resembles the reference curve [39]. In Figure 3 (a)-(c), the prediction results of the proposed model are the predicted curves, and the numerical calculation results are the reference curves.

In Figure 3 (d)-(f), we present the heat maps to depict the curve similarity of predicted and numerically calculated

charge density of negative glow corona under different atmospheric conditions. Here, the numbers in the heat maps represent the value of hausdorff distance. The small distance means a large similarity. If every points of prediction set are close to points of numerical calculation set, the small hausdorff distance value between two kinds of set is obtained. For the new atmospheric condition, the well-trained model takes approximate two seconds to provide the predictions, while the numerical calculation takes 25 minutes. The closeness of machine learning (ML)-based predictions with numerical calculations indicates that the model could instantly predict the charge density with reasonable accuracy.

For the onset condition of positive glow corona, Figure 4 (a)-(c) presents the comparisons of predicted and numerically calculated charge density under different atmospheric conditions (the scale type of the ordinate is log10). Figure 4 (d)-(f) are the heat maps to depict the curve similarity of predicted and numerically calculated charge density of positive glow corona under different atmospheric conditions. The interactive influence of temperature, relative humidity and pressure on the charge density at onset condition has been quantified. As evinced by the high consistency between

the ML predictions and numerical results in Figure 3 and 4, the charge density could be satisfactorily assessed under different atmospheric conditions by the proposed LSTM model with attention mechanisms. The stable corona may appear in glow or streamer mode depending on the electric field and avalanche growth. Similarly, this LSTM model is also applicable to the prediction of charge density at onset condition of streamer corona discharge.

Our results demonstrate that when trained and implemented under the same practices, our proposed model could substantially outperform the traditional time series prediction methods, despite being more complex in our model. In addition, the observations of substantial improvement in prediction performance benefit from the spatial and temporal attentions mechanism when compared with LSTM and RNN. It is attributed to the spatial attention mechanism which could adaptively choose the relevant input series and temporal attention mechanism which could capture relevant long-range encoder hidden states. Note, there may be other ways to potentially improve the predicted performance, such as, using an advanced parameter tuning method or a newer autoencoder architecture. The LSTM model with spatial and temporal attentions is strongly recommended to make prediction based on the spatiotemporal information of avalanche charge. Notably, the capability of attention mechanism to select relevant series in the proposed model does not imply that this model will necessarily perform better for all sequential tasks. In the future, the further robustness work of prediction is recommended, such as, prediction by evaluating on another dataset.

V. CONCLUSION

In summary, a LSTM model with attention mechanisms is introduced for the accelerated prediction of charge density at onset condition of DC corona discharge since the conventionally physical-based onset models are hampered by the computational power and given time. We provide how generated and curated necessary data could be used to train the proposed model. Our results present that the proposed model performs significantly better for the predictive tasks of charge density than ARIMA, ARIMAX, RNN, and LSTM. Notably, our findings also highlight the use of the trained proposed model to successfully and rapidly make predictions with high accuracy for specific atmospheric conditions. Our study motivates future LSTM neural networks with attention mechanisms as a complementary method to numerical calculation and experimental efforts in accelerated prediction for different atmospheric conditions.

REFERENCES

[1] O. M. G. Newman and D. J. Palmer, "Collection of atomic mercury by electrostatic precipitators," *Nature*, vol. 275, no. 5680, pp. 526–527, Oct. 1978.
[2] G.-B. Zhao, S. John, J.-J. Zhang, J. C. Hamann, S. S. Muknahallipatna, S. Legowski, J. F. Ackerman, and M. D. Argyle, "Production of hydrogen and sulfur from hydrogen sulfide in a nonthermal-plasma pulsed corona discharge reactor," *Chem. Eng. Sci.*, vol. 62, no. 8, pp. 2216–2227, Apr. 2007.
[3] M. Rickard, D. Dunn-Rankin, F. Weinberg, and F. Carleton, "Maximizing ion-driven gas flows," *J. Electrostatics*, vol. 64, no. 6, pp. 368–376, Jun. 2006.
[4] I. Levchenko, S. Xu, G. Teel, D. Mariotti, M. L. R. Walker, and M. Keidar, "Recent progress and perspectives of space electric propulsion systems based on smart nanomaterials," *Nature Commun.*, vol. 9, no. 1, pp. 1–19, Dec. 2018.
[5] J. Kieffer, J. P. Matte, M. Chaker, Y. Beaudoin, C. Y. Chien, S. Coe, G. Mourou, J. Dubau, and M. K. Inal, "X-ray-line polarization spectroscopy in laser-produced plasmas," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 48, no. 6, p. 4648, 1993.
[6] P. N. Mikropoulos and V. N. Zagkanas, "Threshold inception conditions for positive DC corona in the coaxial cylindrical electrode arrangement under variable atmospheric conditions," *IEEE Trans. Dielectr., Electr. Insul.*, vol. 22, no. 1, pp. 278–286, Feb. 2015.
[7] A. Eifert, T. Baier, and S. Hardt, "Small onset voltages in negative corona discharges using the edges of gold and aluminum foils as nano-structured electrodes," *Appl. Phys. Lett.*, vol. 103, no. 2, Jul. 2013, Art. no. 023114.
[8] Y. Yi, L. Wang, and Z. Chen, "A comprehensive numerical model of steady-state negative corona for coaxial cylindrical electrode in atmosphere air," *IEEE Trans. Plasma Sci.*, vol. 45, no. 2, pp. 200–207, Feb. 2017.
[9] Z. Bo, K. Yu, G. Lu, S. Cui, S. Mao, and J. Chen, "Vertically oriented graphene sheets grown on metallic wires for greener corona discharges: Lower power consumption and minimized ozone emission," *Energy Environ. Sci.*, vol. 4, no. 7, pp. 2525–2528, 2011.
[10] R. B. Sadeghian and M. Kahrizi, "A novel miniature gas ionization sensor based on freestanding gold nanowires," *Sens. Actuators A, Phys.*, vol. 137, no. 2, pp. 248–255, Jul. 2007.
[11] A. Modi, N. Koratkar, E. Lass, B. Wei, and P. M. Ajayan, "Miniaturized gas ionization sensors using carbon nanotubes," *Nature*, vol. 424, no. 6945, pp. 171–174, Jul. 2003.
[12] G. Naidis, "Conditions for inception of positive corona discharges in air," *J. Phys. D, Appl. Phys.*, vol. 38, no. 13, p. 2211, 2005.
[13] M. Abdel-Salam, M. Nakano, and A. Mizuno, "Corona-induced pressures, potentials, fields and currents in electrostatic precipitator configurations," *J. Phys. D, Appl. Phys.*, vol. 40, no. 7, p. 1919, 2007.
[14] J. Lowke and F. D'Alessandro, "Onset corona fields and electrical breakdown criteria," *J. Phys. D, Appl. Phys.*, vol. 36, no. 21, p. 2673, 2003.
[15] M. Abdel-Salam, A. Turky, and A. Hashem, "The onset voltage of coronas on bare and coated conductors," *J. Phys. D, Appl. Phys.*, vol. 31, no. 19, p. 2550, 1998.
[16] J. F. Hoburg, "Charge density, electric field, and particle charging in electrostatic precipitation with back ionization," *IEEE Trans. Ind. Appl.*, vol. IA-18, no. 6, pp. 666–672, Nov. 1982.
[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, Nov. 2015.
[18] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
[19] Y. Wu, R. Zhao, J. Zhu, F. Chen, M. Xu, G. Li, S. Song, L. Deng, G. Wang, H. Zheng, S. Ma, J. Pei, Y. Zhang, M. Zhao, and L. Shi, "Brain-inspired global-local learning incorporated with neuromorphic computing," *Nature Commun.*, vol. 13, no. 1, pp. 1–14, Dec. 2022.
[20] G. Y. Liu, D. Y. Kong, S. G. Hu, Q. Yu, Z. Liu, T. P. Chen, Y. Yin, S. Hosaka, and Y. Liu, "Smart electronic skin having gesture recognition function by LSTM neural network," *Appl. Phys. Lett.*, vol. 113, no. 8, Aug. 2018, Art. no. 084102.
[21] S. Lu, H. Chai, A. Sahoo, and B. T. Phung, "Condition monitoring based on partial discharge diagnostics using machine learning methods: A comprehensive state-of-the-art review," *IEEE Trans. Dielectr., Electr. Insul.*, vol. 27, no. 6, pp. 1861–1888, Dec. 2020.
[22] A. Aliat, C.-J. Tsai, C.-T. Hung, and J.-S. Wu, "Effect of free electrons on nanoparticle charging in a negative direct current corona charger," *Appl. Phys. Lett.*, vol. 93, no. 15, Oct. 2008, Art. no. 154103.
[23] I. Adamovich, S. D. Baalrud, Anemie Bogaerts, P. J. Bruggeman, M. Cappelli, V. Colombo, U. Czarnetzki, U. Ebert, J. G. Eden, P. Favia, and D. B. Graves, "The 2017 Plasma Roadmap: Low temperature plasma science and technology," *J. Phys. D, Appl. Phys.*, vol. 50, no. 32, 2017, Art. no. 323001.
[24] A. Mesbah and D. B. Graves, "Machine learning for modeling, diagnostics, and control of non-equilibrium plasmas," *J. Phys. D, Appl. Phys.*, vol. 52, no. 30, Jul. 2019, Art. no. 30LT02.
[25] N. O'Connor and S. Daniels, "Passive acoustic diagnostics of an atmospheric pressure linear field jet including analysis in the time-frequency domain," *J. Appl. Phys.*, vol. 110, no. 1, Jul. 2011, Art. no. 013308.

[26] A.-F. Kanta, G. Montavon, M.-P. Planche, and C. Coddet, "Plasma spray process on-line control by artificial intelligence methodology," *Adv. Eng. Mater.*, vol. 9, nos. 1–2, pp. 105–113, Feb. 2007.

[27] A. Bogaerts, X. Tu, J. C. Whitehead, G. Centi, L. Lefferts, O. Guaitella, F. Azzolina-Jury, H.-H. Kim, A. B. Murphy, W. F. Schneider, T. Nozaki, J. C. Hicks, A. Rousseau, F. Thevenet, A. Khacef, and M. Carreon, "The 2020 plasma catalysis roadmap," *J. Phys. D, Appl. Phys.*, vol. 53, no. 44, Oct. 2020, Art. no. 443001.

[28] J. R. Kitchin, "Machine learning in catalysis," *Nature Catal.*, vol. 1, no. 4, pp. 230–232, Apr. 2018.

[29] R. Yang and R. Chen, "Real-time plasma process condition sensing and abnormal process detection," *Sensors*, vol. 10, no. 6, pp. 5703–5723, Jun. 2010.

[30] L. Grajciar, C. J. Heard, A. A. Bondarenko, M. V. Polynski, J. Meeprasert, E. A. Pidko, and P. Nachtigall, "Towards operando computational modeling in heterogeneous catalysis," *Chem. Soc. Rev.*, vol. 47, no. 22, pp. 8307–8348, 2018.

[31] X. Fu, B. Yang, and S. Wang, "Modeling and control of plasma horizontal displacement for HL-2A tokamak based on LSTM," *Fusion Eng. Des.*, vol. 162, Jan. 2021, Art. no. 112107.

[32] K. Zhang, D. Chen, B. Guo, J. Chen, and B. Xiao, "Density limit disruption prediction using a long short-term memory network on EAST," *Plasma Sci. Technol.*, vol. 22, no. 11, Nov. 2020, Art. no. 115602.

[33] G. Farias, E. Fabregas, S. Dormido-Canto, J. Vega, and S. Vergara, "Automatic recognition of anomalous patterns in discharges by recurrent neural networks," *Fusion Eng. Des.*, vol. 154, May 2020, Art. no. 111495.

[34] F. Matos, V. Menkovski, A. Pau, G. Marceca, and F. Jenko, "Plasma confinement mode classification using a sequence-to-sequence neural network with attention," *Nucl. Fusion*, vol. 61, no. 4, Apr. 2021, Art. no. 046019.

[35] N. Qu, Z. Li, J. Zuo, and J. Chen, "Fault detection on insulated overhead conductors based on DWT-LSTM and partial discharge," *IEEE Access*, vol. 8, pp. 87060–87070, 2020.

[36] X. Zhou, X. Wu, P. Ding, X. Li, N. He, G. Zhang, and X. Zhang, "Research on transformer partial discharge UHF pattern recognition based on CNN-LSTM," *Energies*, vol. 13, no. 1, p. 61, Dec. 2019.

[37] M.-T. Nguyen, V.-H. Nguyen, S.-J. Yun, and Y.-H. Kim, "Recurrent neural network for partial discharge diagnosis in gas-insulated switchgear," *Energies*, vol. 11, no. 5, p. 1202, May 2018.

[38] X.-F. Yuan, L. Li, Y. Shardt, Y.-L. Wang, and C.-H. Yang, "Deep learning with spatiotemporal attention-based LSTM for industrial soft sensor model development," *IEEE Trans. Ind. Electron.*, vol. 68, no. 5, pp. 4404–4414, May 2021.

[39] W.-L. Hung and M.-S. Yang, "Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance," *Pattern Recognit. Lett.*, vol. 25, no. 14, pp. 1603–1611, Oct. 2004.

YONG YI (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2017. He performed a Postdoctoral Researcher in electrical engineering at Tsinghua University and The Chinese University of Hong Kong. His research interests include power transmission technology, lightning protection technology, sensor network and big data, dielectric materials and systems, machine learning in power systems, smart dielectric materials, energy conversion and energy collection, electromagnetic environment, and corona discharges.

ZHENGYING CHEN (Member, IEEE) received the M.S. degree in electrical engineering from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2018. He is currently a Senior Engineer with the China Development Bank. His major research interests include high voltage insulation, broadband sensor networks, big data analysis, electrical discharges, dielectric materials and systems, and machine learning in power systems.

RUI LI (Member, IEEE) received the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2019. He was a Visiting Ph.D. student at Harvard University, Cambridge, MA, USA, from September 2017 to September 2018. He was also a Visiting Ph.D. student at the Argonne National Laboratory, Lemont, IL, USA, from November 2016 to March 2017. His research interests include compressed air energy storage, high-performance computation, reinforcement learning, and quantum computing.

· · ·