Forecasting Pavement Performance with a Feature Fusion LSTM-BPNN Model

Yushun Dong
dongyushun887@gmail.com
Beijing University of Posts and Telecommunications
Beijing, China

Yingxia Shao*
shaoyx@bupt.edu.cn
Beijing University of Posts and Telecommunications
Beijing, China

Xiaotong Li
carolineleeton@outlook.com
Beijing University of Posts and Telecommunications
Beijing, China

Sili Li*
decoli27@gmail.com
Research Institute of Highway, Ministry of Transportation
Beijing, China

Lei Quan
563130494@qq.com
Research Institute of Highway, Ministry of Transportation
Beijing, China

Wei Zhang
zhangwei.thu2011@gmail.com
East China Normal University
Shanghai, China

Junping Du
junpingdu@126.com
Beijing University of Posts and Telecommunications
Beijing, China

ABSTRACT

In modern pavement management systems, pavement roughness is an important indicator of pavement performance, and it reflects the smoothness of pavement surface. International Roughness Index (IRI) is the de-facto metric to quantitatively analyze the roughness of pavement surface. The pavement with high IRI not only reduces the lifetime of vehicles, but also raises the risk of car accidents. Accurate prediction of IRI becomes a key task for the pavement management system, and it helps the transportation department refurbish the pavement in time. However, existing models are proposed on top of small datasets, and have poor performance. Besides, they only consider cross-sectional features of the pavements without any time-series information. In order to better capture the latent relationship between the cross-sectional and time-series features, we propose a novel feature fusion LSTM-BPNN model. LSTM-BPNN first learns the cross-sectional and time-series features with two neural networks separately, then it fuses both features via an attention mechanism. Experimental results on a high-quality real-world dataset clearly demonstrate that the new model outperforms existing considerable alternatives.

*Corresponding author.

CIKM '19, November 3–7, 2019, Beijing, China

ACM ISBN 978-1-4503-6976-3/19/11…$15.00
https://doi.org/10.1145/3357384.3357867

CCS CONCEPTS

• **Information systems** → **Data mining**; • **Applied computing** → **Forecasting**; • **Mathematics of computing** → Time series analysis.

KEYWORDS

Pavement Performance Prediction; Feature Fusion; Neural Network; Attention

ACM Reference Format:
Yushun Dong, Yingxia Shao, Xiaotong Li, Sili Li, Lei Quan, Wei Zhang, Junping Du. 2019. Forecasting Pavement Performance with a Feature Fusion LSTM-BPNN Model. In *The 28th ACM International Conference on Information and Knowledge Management (CIKM '19), November 3–7, 2019, Beijing, China.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3357384.3357867

1 INTRODUCTION

Pavement roughness is a vital factor indicating the availability of pavements as well as drivers' comfort. Uneven pavement surface not only affects the running costs of vehicles, such as increasing fuel consumption, reducing driving speed, and extending the traveling time costs, but also increases the risk of car accidents. With the continuous improvement of pavement service and the establishment of modern management system, pavement roughness has become one of the most important indicators of pavement performance. In order to quantitatively analyze the roughness of pavement surface, the International Roughness Index (IRI)[1] is defined as the ratio of total standard body suspension displacement to the distance traveled. Nowadays, IRI has developed into a general indicator of pavement roughness in global.

Prediction model, which provides descriptions and predictions to maintain pavements in serviceable and functional conditions,

[1] https://en.wikipedia.org/w/index.php?title=International_Roughness_Index

Table 1: Comparison of different pavement performance prediction models. Table 2 lists the descriptions of the abbreviations.

References	Targets	Models	Metrics	Features	Data sizes
Hakim et al. [1]	IRI	BPNN	R^2	PSC, CLT, Per	184
Attoh-Okine et al. [3]	IRI	BPNN	Error	PSC, Per	-
Choi et al. [6]	IRI	BPNN	RMSE	PSC, CLT, TRF, Per	117
Gong et al. [11]	Cracking	GBM, XGBoost	MAE	PSC, CLT, TRF, Per	414
Gong et al. [12]	Rutting	BPNN	R^2, RMSE	PSC, CLT, TRF, Per	440
Hossain et al. [16]	IRI	BPNN	RMSE	CLT, TRF, Per,	214
Torre et al. [17]	Δ IRI	BPNN	RMSE	PSC, CLT, TRF, Per	-
Mazari et al. [19]	IRI	BPNN+GEP	R^2, RMSE	TRF, Per	-
Saghafi [26]	Faulting	BPNN	R^2, RMSE	PSC, Per	405
Ziari et al. [33]	IRI	BPNN, GMDH	R^2, RMSE	PSC, CLT, TRF, Per	205
Our paper	**IRI**	**LSTM-BPNN**	**R^2, RMSE**	**PSC, CLT, TRF, Per**	**2243**

is a critical component in pavement management systems, like traffic forecasting [2], IRI prediction, rutting prediction, etc. Among various prediction models, accurate IRI prediction model helps pavement management departments to understand the tendency of pavement performance in time, and thus they are able to refurbish the pavement right away rather than after a longer time interval left for human detection.

Many researchers have been working on IRI prediction models. Traditional experience-based models such as Mechanistic Empirical pavement Design Guide (MEPDG) has been proved to yield worse performance than machine learning models [1]. Up to now, many machine learning models for pavement performance prediction have been developed. One type of the model is tree based model [11, 13]. Another one is Back Propagation Neural Network (BPNN) model [3, 6, 17, 19]. Compared to the experience-based model, machine learning models yield better performance. For one thing, BPNN and tree based models can better capture latent relationships among various features. Besides, compared with traditional models, the inputs of the machine learning models can be easily extended by adding more features to improve the predicting performance.

However, the existing models are mostly built on datasets which contain a small number of data records with simplified pavement features. Therefore, they hardly achieve accurate predictions in real large scale scenarios. Actually, many time-series features are simplified as cross-sectional data, neglecting the varying tendency of the whole series. Table 1 summarizes the statistics of datasets used by previous works. It is easy to figure out that only hundreds of data points are used for learning models, which leads to poor generalization. Second, most of the existing models do not cover the complete feature set (i.e., Per, PSC, CLT, TRF). Even if some models (e.g., Ziari et al. [33]) process the whole feature set; for the climate features, however, they just compute several user-defined statistics (e.g., averages, sums and variances), and fail to capture the deep pattern of climate changes. Consequently, many complex time-series features with latent relationships in real situations are neglected. To the best of our knowledge, none of the previous studies has combined time-series features and cross-sectional features for IRI prediction, in which way specific changes in the series can be captured for the prediction. *In order to obtain an IRI prediction model with high accuracy, the critical challenge is how to effectively fuse these features, including cross-sectional and time series ones.*

In this paper, we first create a high-quality dataset for the IRI prediction with domain experts. The dataset contains 2243 data records and it is about 10 times larger than previous used datasets. On the basis of the dataset, we propose a novel feature fusion model, named LSTM-BPNN, for the IRI prediction task. We make use of Long Short-Term Memory (LSTM) [15] to learn the time-series related features, while using BPNN [14] for the cross-sectional features of different pavements. The two types of features are implicitly fused according to the attention mechanism. The attention method automatically adjusts the weight on time series by considering the influence of the cross-sectional features. We conducted comprehensive experiments on top of the high-quality dataset, and the results demonstrate the effectiveness of the LSTM-BPNN model comparing to different kinds of traditional machine learning models. The new model achieves 0.867 R^2 for the IRI regression.

Our contributions are four-fold:

- We propose a feature fusion LSTM-BPNN model based on neural network to improve the accuracy of IRI prediction.
- We propose an attention mechanism to effectively fuse the cross-sectional and time-series features.
- We create a high-quality and large dataset for the IRI prediction task.
- We conduct empirical studies on a high-quality dataset to demonstrate the advantages of the LSTM-BPNN model.

2 RELATED WORK

In this section, we give a brief review about the prediction models of pavement performance.

Tree-based prediction models. Recently Gong et al. [11] used three different tree models (e.g., gradient boosted model, XGBoost [10], random forest regression) to predict the cracking value, which is an index of pavement performance. Nevertheless, the models above require myriads of pavement surface related features, leading to limited usage because of the complex field measurements. In another work [13], he also proposed a model based on random forest regression for analyzing how IRI is affected by different pavement surface features. The features used as model input still require an

Table 2: The abbreviations of pavement features used in this paper.

Pavement structure and construction (PSC)			
Age	The interval between the pavement first construction and the IRI measurement.	ELE	Elevation of pavement.
LAT	Latitude of pavement.	PM	Material of pavement.
LONGI	Longitude of pavement.	PC	Pavement conditions.
Climate (CLT)			
AWD	Annual wet days.	FI	Annual freeze index.
AS	Annual snowfall.	FT	Annual freeze thaw.
CR	Climate region of pavement.	AAT	Annual average temperature.
AMINT	Annual minimum temperature.	TAP	Total annual precipitation.
AAMAXT	Annual average maximum temperature.	AMAXT	Annual maximum temperature.
AAMINT	Annual average minimum temperature.	AIPD	Annual intense precipitation days.
ADT32	Annual days temperature above 32°C.	ADT0	Annual days temperature below 0°C.
Traffic (TRF)		**Performance (Per)**	
AADTT	Annual average daily truck traffic.	IRI	International Roughness Index.
AADKESAL	Annual equivalent single-axle loads in thousands.	IRI0	The first IRI measured after the latest (re-)construction.

array of field measurements, regardless of the pavement management costs and thus cannot be directly used for IRI prediction when aiming at reducing measuring expenditure.

Single-layered Back Propagation Neural Network. Up to now, many single-layered BPNNs have been proposed. Torre en al. [17] found that the BPNN produced reasonable IRI predictions in several existing flexible sections. Attoh-Okine [3] used real pavement data from Kansas DOT to investigate the effects of learning rate and momentum term on flexible pavement performance prediction. Choi et al. [6] claimed that by using sensitivity analysis neural network could effectively understand the factors controlling overall performance indicators. By comparing different abilities between MEPDG regression and BPNN model on IRI prediction, Abd El-Hakim et al. [1] claimed that BPNN model yields a higher prediction accuracy and a lower bias. Recently, Mazari et al. [19] combined the gene expression programming (GEP) and BPNN, and found that the hybrid method can effectively predict future IRI.

Multiple-layered Back Propagation Neural Network. BPNN with multiple layers offers a number of advantages over the single-layered ones above. Saghafi et al. [26] found a multi-layered BPNN structure of 4 layers outperforms traditional method in faulting prediction. In 2014, Ceylan et al. [4] suggested that the BPNN predicts the future of pavement condition with satisfactory accuracy in both short and long term. Then Ziari et al. [33] developed a 3-layered BPNN and the group method of data handling model (GMDH) for the IRI prediction. Recently, Hossain et al. [16] also designed a Deep Neural Network model for the prediction of IRI by using various statistical characteristics on traffic and climate. However, the proposed models were only tested on a relatively small dataset, which is summarized in Table 1. In addition, Gong et al. [12] introduced a deep learning neural network model for rutting prediction with 3 hidden layers, but the climate time-series are still used as cross-sectional features. Multiple-layered BPNN has been used in other fields as well. Covington et al. [7] found that deep collaborative filtering model was able to effectively assimilate many signals and modeled their interaction with layers of depth, outperforming previous matrix factorization approaches used at YouTube [28] . In [18], Liu et al. discussed myriads of widely-used deep learning architectures and their practical applications.

Time-series Prediction. For time-series prediction, researchers have also make thoroughly explorations. In [30], the author proposed Long- and Short-term Time-series network (LSTNet) with temporal attention layer to address the open challenge of multivariate time series forecasting. In [27], the author proposed a new unified model, Grid-Embedding based Multi-task Learning (GEML), to satisfy the need of predicting the number of passenger demands from one region to another in advance. To capture temporal trends of passenger demands, the author designed a Multi-task Learning network resorting to Long Short-Term Memory Recurrent Networks (LSTMs) in order to assist the main task to capture stronger intrinsic temporal patterns, and the proposed model was then proved outperforms the baselines. Apart from that, in [5], the author proposed a new attentional intention-aware recommender (AIR) systems to predict category-wise future user intention and collectively exploit the rich heterogeneous user interaction behaviors by extracting sequential patterns from the user-item interaction data; in [30], the author used point process for modeling event, arguing that the background of many point processes can be treated as a time series. The author took an RNN perspective to point process, and models its background and history effect. In this paper, we will introduce deep neural network to model time-series features of pavements.

3 PRELIMINARY

3.1 Problem Definition

Here we give the definition of the IRI prediction problem. Assume the pavement record set is denoted as $\mathcal{R} = \{\mathcal{R}^{tr}, \mathcal{R}^{te}\}$, where $\mathcal{R}^{tr}$ is used for training and $\mathcal{R}^{te}$ for testing. Each record $r \in \mathcal{R}$ is expressed to be $r = \{p, X_p, y_p\}$, where p denotes the pavement in the record, X_p denotes the cross-sectional as well as time-series

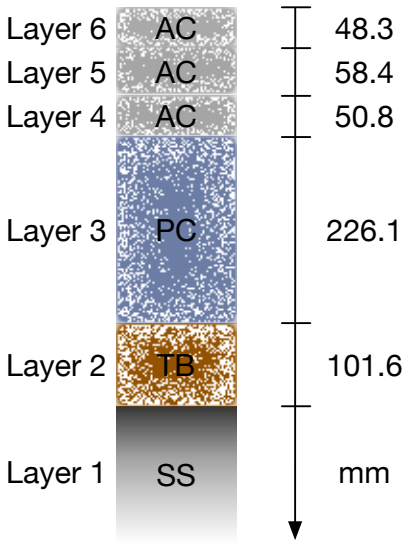


Figure 1: Structure of a pavement in Michigan with six layers.

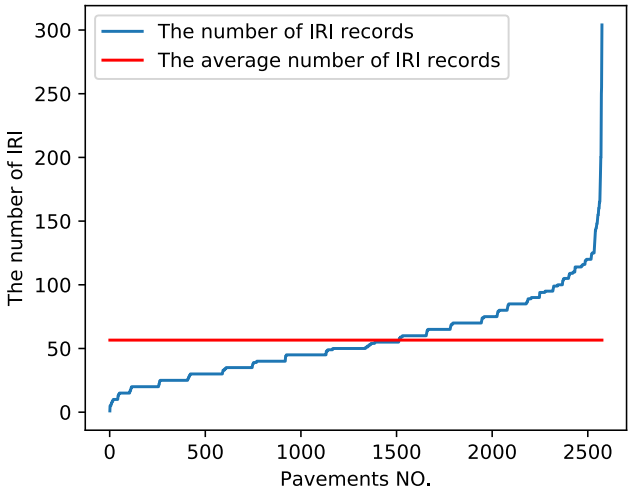


Figure 2: The number of IRI records with different pavements in LTPP database. Best viewed in color.

features of the pavements, and y_p represents the value of IRI, which is the target feature of the prediction model.

Based on the denotations above, we formally define the problem as below,

Definition 1. ***(Pavement IRI Prediction)*** *Given a training set $\mathcal{R}^{tr}$ of pavements, the goal is to learn a model*

$$f : X_p \rightarrow y_p$$

for each record $r \in R^{tr}$, and further utilize the model f to make IRI predictions for records in the test set R^{te}.

3.2 LTPP Database

The Long-Term Pavement Performance (LTPP) program [23], which is initiated in 1987, is envisioned as a comprehensive program to satisfy a wide range of pavement information needs [24]. Currently, the LTPP acquires the largest road performance database [21, 25]. A large number of pavement performance related researches [1, 6, 8, 16, 17, 19, 20, 31, 33] use the database to organize their datasets. Overall, there are four kinds of pavement features, namely pavement structure and construction (PSC), climate (CLT), traffic (TRF) and performance (Per). Table 2 summarizes the descriptions of the important features from the four kinds above, and please refer to Section 4 for detailed introduction.

3.3 Pavement Structure

The pavement structure usually consists of 2 to 13 layers, among which some layers use the same material. Fig. 1 illustrates the layer structure pattern of a pavement in Michigan. From top to bottom, first, several asphalt concrete (AC) layers cover on the very top of the pavement surface, which are denoted by layer 4, 5 and 6. Then portland cement concrete (PC) and treated bound (TB) base are followed beneath. At the very bottom is the subgrade (SS) layer (layer 1), which is the base of the whole pavement. In the LTPP database, the data of pavement structure is recorded layer-by-layer. Note that different pavements have different combinations of layers.

4 DATA ORGANIZATION

We use LTPP database as the data source to create high-quality datasets for model design. In the following paragraphs, we first introduce the problems of raw LTPP database for IRI prediction. Then we describe the approaches of extracting high-quality features, including cross-sectional and time-series ones.

4.1 LTPP Database for IRI prediction

Although LTPP database has been in service for around thirty years and accumulates a large amount of data, the raw data have the following problems preventing researchers from modeling IRI accurately.

Sparse observations. For most pavements, IRIs are only measured once a year or every two years. Such low frequency of measurement causes the data sparsity problem. Fig. 2 visualizes the distribution of the number of IRI records among different pavements. In average, each pavement only has about 50 IRI records. Even for the oldest pavement (e.g., 70 years), the number of IRI data is still relatively small (e.g., 300 rows). The sparsity problem hinders previous works [1, 6, 16, 17, 19, 33], which used a few number of roads, having a good generalization.

Various temporal and cross-sectional data. There are at least two types of data which influence the IRI of pavements. One is cross-sectional features, such as traffic, location and structure. The other is time-series features, like climate. Being observed automatically by electronic equipment, climate features are recorded more regularly and have fewer data missing than cross-sectional data in LTPP database [21] . For example, air temperature and precipitation are collected at every 15 minutes, and these data are accumulated into hourly, daily, and monthly statistics. Previous studies [1, 16, 17, 33]

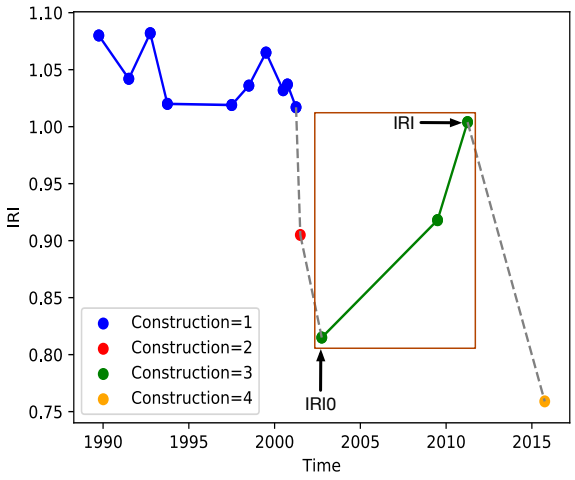


Figure 3: An IRI tendency of pavement 25-1004 in Massachusetts. The reconstruction events are denoted with different colors. Best viewed in color.

have demonstrated that both features affect the predictions of pavement performance. However, there is no model to reasonably fuse these two kinds of features for capturing more latent relationships.

Human factors. Human factors are introduced to the IRI tendency by pavement reconstructions. After a reconstruction, IRI drops obviously due to the pavement surface quality improvement, and then it will rise gradually with time going on. Fig. 3 illustrates such trend with pavement 25-1004 in Massachusetts. Furthermore, the reconstruction operation breaks the whole IRI sequence of a single pavement into several short series where human factors are excluded. Therefore, the valid IRI sequences are short and it is hard to model its time-dependent pattern directly.

Summary. In this paper, to overcome the problems above, we mix all pavements together to guarantee that we have enough data for modeling, and develop a feature fusion model based on deep neural network by utilizing both cross-sectional and time-series data. In the next two subsections, we describe the detailed data pre-processes for feature extraction.

4.2 Cross-sectional features

We focus on pavement performance, pavement structure and construction, and traffic to generate useful cross-sectional features. Table 3 lists the statistics of the cross-sectional features used in this paper.

Pavement performance. In this work, we only use IRI related features at the aspect of pavement performance. Due to human factor, the IRI sequence is short, and we model IRI as a cross-sectional feature instead of time-series feature. For each IRI observation date, we extract two IRI values: one is the current IRI value, and the other is the first IRI measurement in this pavement construction, denoted by IRI0. Taking the construction=3 of pavement 25-1004 (green line in Fig. 3) as an example, IRI0 is the first point denoted in the graph, and one IRI value is the end point of the line. Besides, to ensure

Table 3: Cross-sectional features statistics. N/A is Not Applicable.

Feature	Range	Average	Unit
IRI	0.3460~4.473	1.478	m/km
IRI0	0.333~4.447	1.404	m/km
ELE	-11~2331	486.428	m
LAT (abs)	18.33~62.41	38.905	degree
LONGI (abs)	52.87~156.67	96.487	degree
AGE	307~29331	7163.227	days
AADTT	7~9353	846.295	N/A
AADKESAL	1~5214	404.768	time

Table 4: Climate features statistics

Feature	Range	Average	Unit
AAT	-10.9~26.6	12.402	°C
FI	0.0~2370.0	326.420	°C ∗ days
FT	0~209	78.709	days
AAMAXT	-6.4~32.6	18.875	°C
AAMINT	-15.4~21.7	5.869	°C
AMAXT	6.8~51.6	36.522	°C
AMINT	-43.8~18.4	-17.537	°C
ADT32	0~197	40.785	days
ADT0	0~254	103.898	days
TAP	10.5~2800.8	840.139	mm
AS	0.0~4411.0	672.645	time
AIPD	0~73	19.167	days
AWD	6~267	127.402	days

that the difference between IRI and IRI0 is influenced by time-series features for a sufficient long period, the extracted IRI records are measured at least two years after each pavement reconstruction.

Pavement structure and construction. For structure and construction related features, we use construction number of each observation to denote which time period the pavement structure records should be extracted from, i.e., given a pavement, for each distinct construction number, the dates between the first IRI observation and the last IRI observation are extracted. Apart from that, we also need basic pavement features such as the latitude (LAT), longitude (LONGI), elevation (ELE) and Age, because we use all pavement data to create the prediction model, and the model needs the basic features to distinguish each pavement. Consequently, we extract all pavement location related data as shown in Table 2. Here pavement age of an IRI observation is the days from the date pavement opened to the date IRI measured. Lastly, we also extract features about the pavement layer structure, namely layer number, layer thickness and layer material. One-hot encode is adopted for the representation of the structure of the pavements.

Traffic. We extract two types of traffic features as shown in Table 2, and they are annual equivalent single-axle loads (AADKESAL) and annual average daily truck traffic (AADTT). AADKESAL represents a converting result of damage from wheel loads of various magnitudes and repetitions ("mixed traffic") to damage from an equivalent number of "standard" or "equivalent" loads [21], and

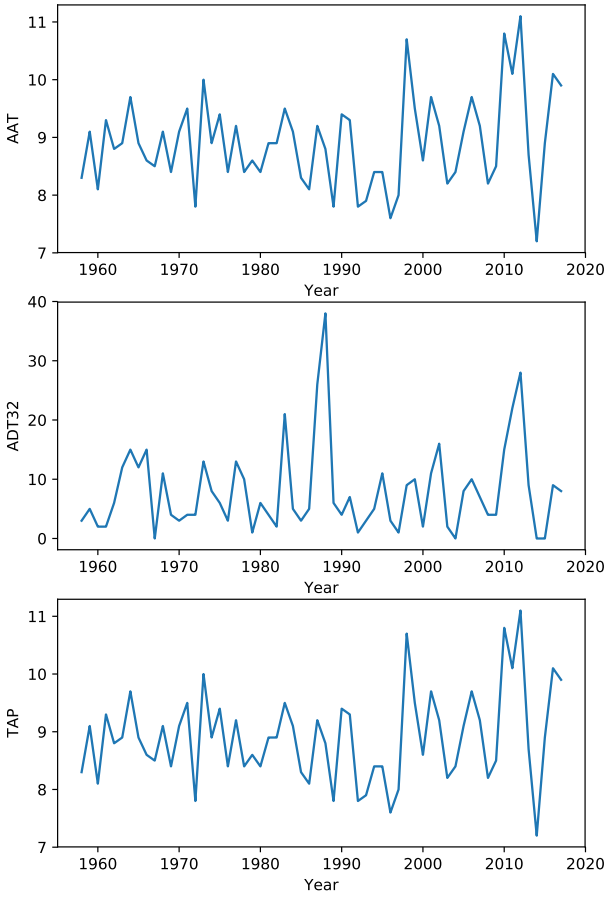


Figure 4: Different time-series features of pavement 26-116 in Michigan

it represents how much the pavement is worn and crushed. Besides, AADTT plays a key role in influencing IRI tendency since the tremendous weights of the trucks heavily affect the conditions of the pavement surfaces. For the missing values of AADKESAL and AADTT, we fill in them with the latest data near the records considering the traffic of each pavement to be relatively stable.

4.3 Time-series Features

The climate feature is one of the most influential features affecting IRI and we try to capture the latent relationship between IRI tendency and annual climate changes. Fig. 4 shows the complexity patterns of AAT, ADT32 and TAP of pavement 26-116 in Michigan.

In this work, we mainly extract annual temperature and precipitation related features as our time-series features. There are nine different annual temperature features, and they are AMINT, AAMINT, AMAXT, AAMAXT, AAT, FT, FI, ADT32 and ADT0. The annual precipitation includes AWD, TAP, AIPD and AS. Because the time period between IRI measurement and IRI0 is varying, we need to create time-series features with varying length. However, our model (introduced in Section 5) requires fixed-length input.

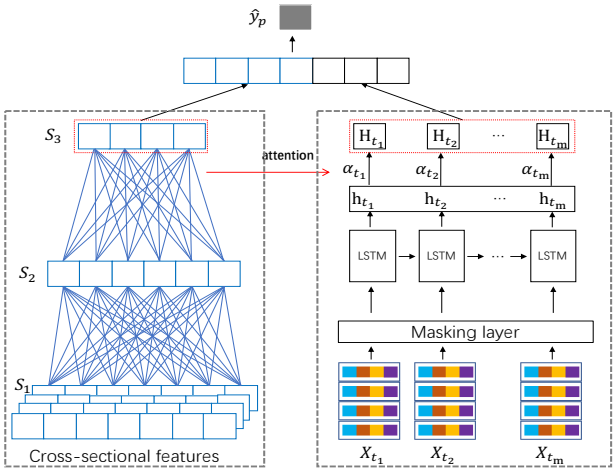


Figure 5: The architecture of feature fusion LSTM-BPNN Model

Therefore, we first extract all climate features for 20 years, ensuring the time is fully covered from the lastest reconstruction to the IRI measurements in LTPP database. Then we delete climate data earlier than IRI0 date and use a magic number (e.g., -100) to fill in the deleted value in climate series. Table 4 lists the statistics of time-series features used in this paper.

5 FEATURE FUSION LSTM-BPNN MODEL

5.1 Model Specification

Fig. 5 describes the architecture of our feature fusion LSTM-BPNN model. It mainly consists of two components: 1) the left part in the figure illustrate a BPNN, which accepts the basic and cross-sectional features of different pavements as inputs and learns corresponding hidden representations, and 2) the right part presents a recurrent neural network for the modeling of annual climate time-series features with various length. We apply the attention mechanism to learn time-series features by implicitly fusing the cross-sectional features. The two components work in a manner of affecting each other mutually. The final output representation of LSTM-BPNN concatenates the output of BPNN and LSTM.

BPNN model for cross-sectional features. To model the cross-sectional features of pavements, we apply neural network, BPNN model, for its high expressivity. In this way, our model will have better capability to capture important features on pavement itself.

Consider the state of the jth neuron of layer $l-1$ to be β_j^{l-1}, we define the following equation:

$$\beta_j^l = relu\left(\sum_k w_{jk}^l \beta_k^{l-1} + b_j\right), \tag{1}$$

where $w_j k^l$ is a weight matrix and b_j is a bias vector. In our model, we set the layer number of BPNN component to be 3, namely S_1, S_2, S_3. S_1 is the input layer of the cross-sectional features. S_2 is the hidden layer and S_3 is the top layer for the final matrices concatenating with the output of LSTM component.

LSTM for climate time-series features. Long short-term memory (LSTM) has been successfully applied to model time-series features [22]. Therefore, we adopt it for encoding the climate time-series features and better capturing the impacts of climate changes over time on pavement performance. Actually, LSTM contributes to the detection of deleterious time intervals in sequences for each pavement, in which way the fluctuations can be captured comparing to the traditional methods which use only statistic characteristics to represent impacts.

The LSTM part mainly consists of three components: the forget gate, the input gate and the output gate. Consider the input X_{t_c} at time t_c, the previous hidden layer state $h_{t_{c-1}}$, and the forget gate state f_{t_c} at time t_c, we have:

$$f_{t_c} = sigmoid\left(W_f \cdot \left[h_{t_{c-1}}, X_{t_c} + b_f\right]\right). \tag{2}$$

Similarly, consider the state of the output gate to be C_{t_c} and the two middle states $\widetilde{C}_{t_c}$ and i_{t_c}, we have:

$$i_{t_c} = sigmoid\left(W_i \cdot \left[h_{t_{c-1}}, X_{t_c}\right] + b_i\right), \tag{3}$$

$$\widetilde{C}_{t_c} = \tanh\left(W_C \cdot \left[h_{t_{c-1}}, X_{t_c}\right] + b_C\right), \tag{4}$$

$$C_{t_c} = f_{t_c} * C_{t_{c-1}} + i_{t_c} * \widetilde{C}_{t_c}. \tag{5}$$

Finally, with the output gate state $h_{t_{c-1}}$, we have:

$$o_{t_c} = sigmoid\left(W_o \cdot \left[h_{t_{c-1}}, X_{t_c} + b_o\right]\right), \tag{6}$$

$$h_{t_c} = o_{t_c} * \tanh\left(C_{t_c}\right), \tag{7}$$

where W_f, W_i, W_o and W_C are weight matrices, and b_f, b_i, b_C and b_o are weight vectors.

Masking layer. Considering that the length of time-series features with valid values is variable, we add a masking layer before the input of LSTM component. The masking layer will automatically filter invalid features (e.g., the magic number -100) and process features with varying length.

Attention-based Feature Fusion. For IRI prediction, it has been clearly that different climate changes affect the performance of pavements. In our model, we are seeking to take more details into consideration. Specifically, we would like to model how the cross-sectional features of the pavements (e.g., the structure of the pavement and traffic related features) and the different changes of climate influence each other. Here we apply attention mechanism, since it is able to capture more latent relationship [29]. By using attention method in LSTM-BPNN model, cross-sectional features affect the weights of the climate features in a year-by-year manner, and the weight α will be automatically aggravated or reduced by the learning process. Before introducing the novel attention mechanism, we first define S_3 and S_2 to be the intermediate and top layers' outputs of the BPNN, respectively, and assume each layer of the network is associated with nonlinear activation functions (e.g., rectified linear unit). Then we propose a novel method to compute the overall impact of each annual climate at time t as α_{t_i} with pavements features being taken into consideration:

$$\begin{bmatrix} \alpha_{t_1} \\ \vdots \\ \alpha_{t_m} \end{bmatrix} = f\left(W_\alpha \cdot \tanh\left(W_h \mathbf{h_t} + W_s S_2 + b_\alpha\right)\right), \tag{8}$$

where W_α, W_h and W_s are weight matrices, $\mathbf{h_t}$ is the matrix of h_{t_i}, and b_α is a bias vector. In addition, function f is the output activation of the attention, and we can set f to be sigmoid or softmax interchangeable. With such attention method, the final output is:

$$H_t = \begin{bmatrix} \alpha_{t_1} \\ \vdots \\ \alpha_{t_m} \end{bmatrix} \cdot \begin{bmatrix} h_{t_1} & \dots & h_{t_m} \end{bmatrix}. \tag{9}$$

IRI regression. After getting the output of the two components to be S_3 and H_t respectively, the model computes the final regression by concatenating H_t and S_3 together. Supposing the target is denoted as $\hat{y}_p$, we define the output of the regression with linear activation as follows:

$$\hat{y}_p = W_y \cdot [H_t; S_3] + b_y, \tag{10}$$

where W_y and b_y are learnable weight matrices and bias vector respectively. We define the loss function of LSTM-BPNN model as the mean square error between the real regression target y_p and the predicated target $\hat{y}_p$.

6 EXPERIMENTS

6.1 Experimental Settings

We implemented our model with Keras library. During the training, early stopping strategy and Adam optimizer with exponential learning rate decay are applied. The parameter of decay steps is set to be 10. For hyperparamters, we use grid search to find the best settings. Table 5 lists the detailed space of the hyperparameters in our model.

Table 5: Hyperparameter Space. The bold text indicates the best parameters.

Hyperparameter	value
BPNN hidden layer activation	[**relu**, tanh]
Attention hidden layer activation	[**tanh**, relu]
Attention output layer activation	[**sigmoid**, softmax]
LSTM hidden states	**13** [8, 64]
Concatenate layer LSTM neurons	**100** [8, 512]
Concatenate layer ANN neurons	**300** [8, 512]
Output layer activation	linear
Epoch	**65** [8, 512]
Batch size	**512** [8, 1024]
Learning rate decay steps	**10** [5, 20]

Baselines. We choose three traditional regression models which are Linear Regression (LR) [32], Gradient Boosting Decision Tree (GBDT), eXtreme Gradient Boosting regression (XGBR), and a pure BPNN model for comparison. Traditional and neuron network models are implemented with Sklearn and Keras library, respectively. Tree based models, namely GBDT and XGBR, are carefully tuned, and the number of boosted trees are set to be 400. BPNN model has

Table 6: Evaluation results of only modeling cross-sectional features.

Model	R^2	RMSE
LR	0.735	0.352
GBDT	0.758	0.336
XGBR	0.767	0.328
BPNN	0.769	0.322

Table 7: Evaluation results of modeling on time-series features and IRI0.

Model	R^2	RMSE
LR (Agg.)	0.732	0.352
LSTM	0.768	0.335
GBDT (Agg.)	0.770	0.327
XGBR (Agg.)	0.771	0.321

1 hidden layer. The batch size and training epoch are set to be 128 and 512, and the output layer of BPNN is linear for regression.

Dataset. We validated our new prediction model using a real-world dataset from LTPP. According to the data organization introduced in Section 4, we obtain a dataset which contains 2243 records and covers 1406 pavements in north America. The size of our dataset is much larger than the ones in previous works. Besides, because different pavements have different layer structure, the features of the pavement structure are sparse, and we use PCA [9] to reduce the dimensions. Specifically, the features of the pavement structure are reduced to 5 dimensions.

Metrics. The evaluation metrics adopted in the experiments are mean square error (RMSE) and coefficient of determination (R^2), which are commonly used in regression tasks. The formulas of the metrics are shown below

$$R^2 = 1 - \frac{\sum (y_i - f_i)^2}{\sum \left(y_i - \frac{1}{n}\sum_{i=1}^{n} y_i\right)^2}, \tag{11}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - f_i)^2}{n}}, \tag{12}$$

where y_i and f_i denote the real and predicted value, respectively. Furthermore, we ran each model in 5-fold cross-validation manner, and reported the average performance.

6.2 The Influence of Cross-Sectional Features

To demonstrate the effectiveness of cross-sectional features for IRI prediction, we first tested different models by only using the cross-sectional features. Note that in this experiment, time-series features are not considered, therefore, we only use the BPNN component in LSTM-BPNN model for comparison. Table 6 shows the results of predictions. All models achieve an R^2 greater than 0.7, indicating the effectiveness of all chosen models. The worst comes from the LR model, which is 0.735. The best performance comes from BPNN regression model, which is 0.769, indicating that neural network has a strong ability for capturing non-linear and complex latent information.

Table 8: Evaluation results of modeling all features.

Model	R^2	RMSE
LR (All)	0.738	0.350
BPNN (All)	0.774	0.353
GBDT (All)	0.781	0.320
XGBR (All)	0.789	0.316
LSTM-BPNN (Softmax)	**0.855**	**0.258**
LSTM-BPNN (Sigmoid)	**0.867**	**0.242**

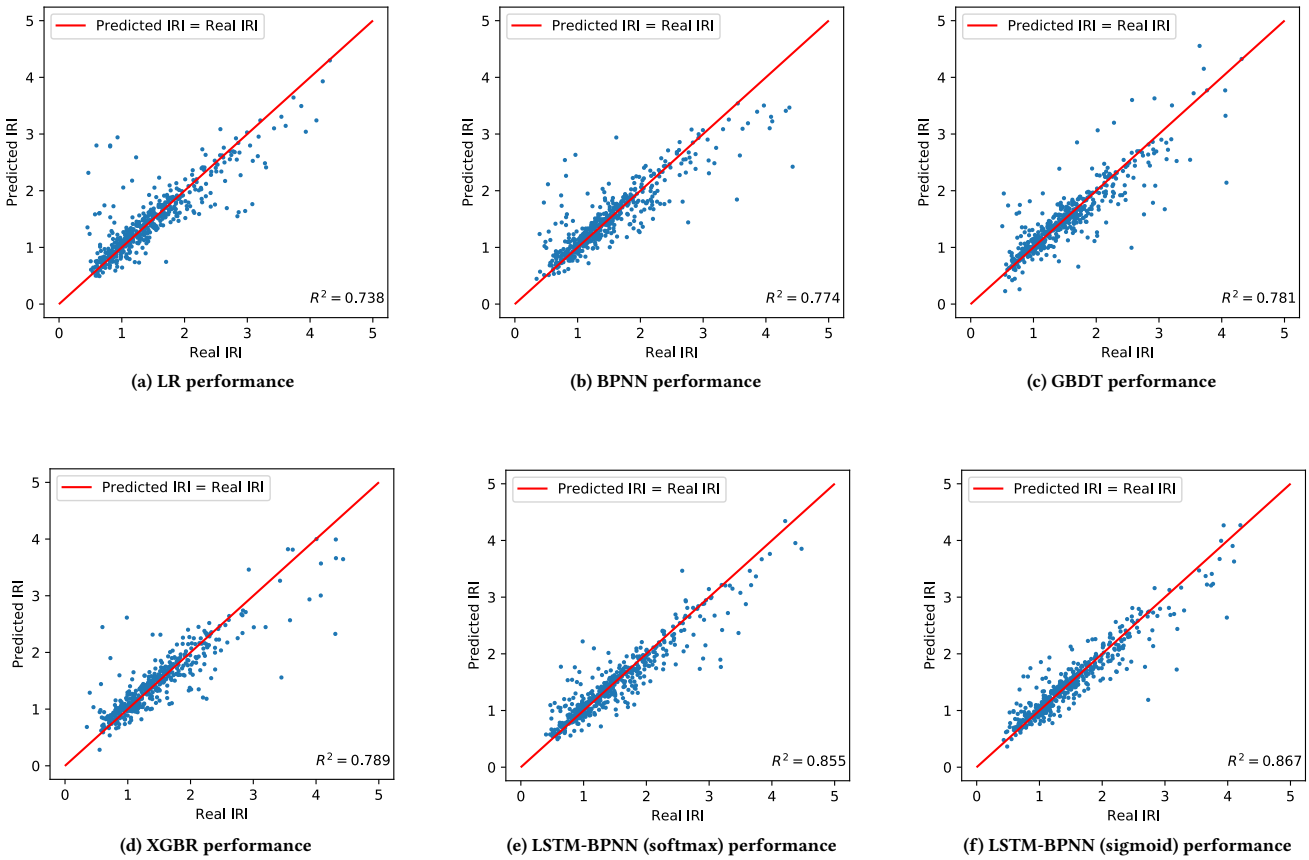
6.3 The Influence of Time-series Features

We also conducted experiments to showcase the effectiveness of time-series features. First in this experiment, we also preserve one cross-sectional feature IRI0 for model training. The reason is that the IRI value is highly related to IRI0. From Table 3, it is easy to figure out that the range of IRI and IRI0 are close, which implies the absolute values of IRI and IRI0 with the same construction number are close. Second, the traditional models cannot process the raw time-series features directly, we compute the aggregated statistics (e.g., average and variance) of each time-series feature instead. These models are denoted with "(Agg.)" suffix. For LSTM-BPNN model, we use only IRI0 for the input of BPNN component and time-series features for the input of LSTM component. As shown in Table 7, among models using the aggregated features of time-series climate features, LR model gives the worst result of R^2 0.732, while XGBR achieves the best and the R^2 is 0.771. Note that LSTM obtains an R^2 of 0.768, which is close to the best one. This implies that directly modeling the time-series features with sequential neural network is a comparative solution.

6.4 The Performance of LSTM-BPNN model

Finally, we compared LSTM-BPNN model with the alternatives using all cross-sectional features and the time series for IRI prediction. Table 8 shows the results. First, more key features clearly improve the prediction accuracy of these models. Among the baselines, the worst performance of 0.738 comes from LR. The best performance comes from XGBR model, which is 0.789. Our LSTM-BPNN model achieves R^2 to be 0.867, which is far better than XGBR model, and also obtains the lowest RMSE to be 0.242. It is also a great improvement of our model comparing to BPNN. Fig. 6 visualizes all the regression results of the six models above, and from the figure, we can also find that LSTM-BPNN achieves better performance of regression compared to other baselines. These experimental results demonstrate that the LSTM model effectively captures the latent relationship between different kinds of climate changes and IRI tendency via the attention mechanism from BPNN, demonstrating that when time-series features is properly taken advantage of, more latent information can be extracted by the model, and finally contributes to more accurate predictions of future IRI.

Attention Activation: Sigmoid vs. Softmax. We also compared our model by setting the attention output activation to be sigmoid and softmax, respectively. As shown in Table 8 and Fig. 6, sigmoid activation gives better performance than softmax. A reasonable explanation is that the impact of annual climate change for pavements is independent every year, which means how the



(a) LR performance (b) BPNN performance (c) GBDT performance

(d) XGBR performance (e) LSTM-BPNN (softmax) performance (f) LSTM-BPNN (sigmoid) performance

Figure 6: The regression results of six models with all features.

IRI will be influenced in the future has nothing to do with how it was influenced before, and vice versa. As a consequence, it is better to compute each attention weight α_{ti} individually without normalization. In other words, sigmoid activation is more suitable for the IRI prediction under time-series attention circumstances.

7 CONCLUSION

IRI prediction is a core task for a pavement management system. In this paper, we proposed a deep learning model for the task, named LSTM-BPNN. The new model first applies BPNN and LSTM to encode cross-sectional and time-series features respectively, and then uses an attention mechanism to capture the latent relationships between them. We also organized a high-quality and large dataset for the pavement IRI prediction from LTPP database. The experimental results on the dataset clearly demonstrate the effectiveness of the LSTM-BPNN for IRI prediction. Besides IRI, there are many other pavement performance indicators, like rutting, cracking, faulting, etc. In the future, we will extend our model to predict these indicators.

ACKNOWLEDGEMENTS

This work is supported by the National Key R&D Program of China (No. 2017YFC0840200), and National Natural Science Foundation of China (No. 61702015, 61532006). The authors thank the anonymous reviewers for their most helpful remarks.

REFERENCES

[1] Ragaa Abd El-Hakim and Sherif El-Badawy. 2013. International roughness index prediction for rigid pavements: an artificial neural network application. In *Advanced Materials Research*, Vol. 723. Trans Tech Publications Ltd, 854–860.

[2] Amr Abdullatif, Francesco Masulli, and Stefano Rovetta. 2017. Tracking Time Evolving Data Streams for Short-Term Traffic Forecasting. *Data Science and Engineering* 2, 3 (01 Sep 2017), 210–223.

[3] Nii O Attoh-Okine. 1999. Analysis of learning rate and momentum term in back propagation neural network algorithm trained to predict pavement performance. *Advances in Engineering Software* 30, 4 (1999), 291–302.

[4] Halil Ceylan, Mustafa Birkan Bayrak, and Kasthurirangan Gopalakrishnan. 2014. Neural networks applications in pavement engineering: A recent survey. *International Journal of Pavement Research and Technology* 7, 6 (2014), 434–444.

[5] Tong Chen, Hongzhi Yin, Hongxu Chen, Rui Yan, Quoc Viet Hung Nguyen, and Xue Li. 2019. AIR: Attentional Intention-Aware Recommender Systems. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. IEEE, 304–315.

[6] Jae-ho Choi, Teresa M Adams, and Hussain U Bahia. 2004. Pavement Roughness Modeling Using Back-Propagation Neural Networks. *Computer-Aided Civil and Infrastructure Engineering* 19, 4 (2004), 295–303.

[7] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. ACM, 191–198.

[8] Qiao Dong and Baoshan Huang. 2011. Evaluation of effectiveness and cost-effectiveness of asphalt pavement rehabilitations utilizing LTPP data. *Journal of Transportation Engineering* 138, 6 (2011), 681–689.

[9] Imola K Fodor. 2002. *A survey of dimension reduction techniques*. Technical Report. Lawrence Livermore National Lab., CA (US).

[10] Fangcheng Fu, Jiawei Jiang, Yingxia Shao, and Bin Cui. 2019. An Experimental Evaluation of Large Scale GBDT Systems. *PVLDB* (2019).

[11] Hongren Gong, Yiren Sun, and Baoshan Huang. 2019. Gradient Boosted Models for Enhancing Fatigue Cracking Prediction in Mechanistic-Empirical Pavement Design Guide. *Journal of Transportation Engineering, Part B: Pavements* 145, 2 (2019), 04019014.

[12] Hongren Gong, Yiren Sun, Zijun Mei, and Baoshan Huang. 2018. Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks. *Construction and Building Materials* 190 (2018), 710–718.

[13] Hongren Gong, Yiren Sun, Xiang Shu, and Baoshan Huang. 2018. Use of random forests regression for predicting IRI of asphalt pavements. *Construction and Building Materials* 189 (2018), 890–897.

[14] Simon Haykin. 1998. *Neural Networks: A Comprehensive Foundation* (2nd ed.). Prentice Hall PTR, Upper Saddle River, NJ, USA.

[15] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[16] MI Hossain, LSP Gopisetti, and MS Miah. 2018. International roughness index prediction of flexible pavements using Neural Networks. *Journal of Transportation Engineering, Part B: Pavements* 145, 1 (2018), 04018058.

[17] Francesca La Torre, Lorenzo Domenichini, and Michael I Darter. 1998. Roughness prediction model based on the artificial neural network approach. In *Fourth International Conference on Managing Pavements*, Vol. 2.

[18] Weibo Liu, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu, and Fuad E Alsaadi. 2017. A survey of deep neural network architectures and their applications. *Neurocomputing* 234 (2017), 11–26.

[19] Mehran Mazari and Daniel D Rodriguez. 2016. Prediction of pavement roughness using a hybrid gene expression programming-neural network technique. *Journal of Traffic and Transportation Engineering (English Edition)* 3, 5 (2016), 448–455.

[20] Alaeddin Mohseni. 1998. *LTPP seasonal asphalt concrete (AC) pavement temperature models*. Technical Report.

[21] Jean Nehme. 2017. *About Long-Term Pavement Performance*. Federal Highway Administration. Retrieved October 22.

[22] David MQ Nelson, Adriano CM Pereira, and Renato A de Oliveira. 2017. Stock market's price movement prediction with LSTM neural networks. In *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1419–1426.

[23] United States Department of Transportation. 2019. Long-Term Pavement Performance. https://highways.dot.gov/long-term-infrastructure-performance/ltpp/long-term-pavement-performance. (2019). [Online; accessed 9-May-2019].

[24] RW Perera and Starr D Kohn. 2001. *LTPP data analysis: Factors affecting pavement smoothness*. Transportation Research Board, National Research Council Washington, DC, USA.

[25] Robert Raab. 2017. *Long-Term Pavement Performance Studies*. Transportation Research Board. Retrieved October 22.

[26] Behrooz Saghafi, Abolfazl Hassani, Roohollah Noori, and Marcelo G Bustos. 2009. Artificial neural networks and regression analysis for predicting faulting in jointed concrete pavements considering base condition. *International Journal of Pavement Research and Technology* 2, 1 (2009), 20–25.

[27] Yuandong Wang, Hongzhi Yin, Hongxu Chen, Tianyu Wo, Jie Xu, and Kai Zheng. 2019. Origin-destination matrix prediction via graph convolution: a new perspective of passenger demand modeling. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 1227–1235.

[28] Jason Weston, Samy Bengio, and Nicolas Usunier. 2011. Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

[29] Bin Xia, Yun Li, Qianmu Li, and Tao Li. 2017. Attention-based recurrent neural network for location recommendation. In *2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*. IEEE, 1–6.

[30] Shuai Xiao, Junchi Yan, Xiaokang Yang, Hongyuan Zha, and Stephen M Chu. 2017. Modeling the intensity function of point process via recurrent neural networks. In *Thirty-First AAAI Conference on Artificial Intelligence*.

[31] Amber Yau, Harold L Von Quintus, et al. 2002. *Study of LTPP laboratory resilient modulus test data and response characteristics*. Technical Report. Turner-Fairbank Highway Research Center.

[32] Lele Yu, Lingyu Wang, Yingxia Shao, Long Guo, and Bin Cui. 2018. GLM+: An Efficient System for Generalized Linear Models. In *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*. 293–300.

[33] Hasan Ziari, Jafar Sobhani, Jalal Ayoubinejad, and Timo Hartmann. 2016. Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *International journal of pavement engineering* 17, 9 (2016), 776–788.