Hindawi
Modelling and Simulation in Engineering
Volume 2020, Article ID 8909747, 14 pages
https://doi.org/10.1155/2020/8909747

Research Article

Service-Life Study of Polycarbonate Outdoors Using Python with Incomplete Data

Jiangfeng An,[1,2] **Duncheng Peng**,[1,2] **Xuejie Zhou**,[2] **Jun Wu**,[2] **and Penghua Zheng**[2]

[1]*Wuhan Research Institute of Materials Protection Co., Ltd, 126 Bao Feng Erlu, Wuhan, Hubei 430030, China*
[2]*State Key Laboratory of Special Surface Protection Materials and Applied Technology, 126 Bao Feng Erlu, Wuhan, Hubei 430030, China*

Correspondence should be addressed to Jiangfeng An; 33104533@qq.com

Received 5 December 2019; Accepted 23 April 2020; Published 26 May 2020

Academic Editor: Zhiping Qiu

The deterioration of polycarbonate (PC) depends on various environmental factors. Meanwhile, the complexity of the related weathering processes inhibits the prediction of service life based on the environmental factors. To elucidate the nonlinear correlation between PC weathering and the environmental factors, three-year-long natural weathering tests were conducted at eight experimental stations in China. The relationship between tensile-property data of PC and environmental and pollutant data is analyzed by extra-trees and multilayer perceptron networks implemented in Python. The results indicated that (1) the degradation of PC tensile properties is mainly affected by the experimental period (76.37%), whilst the effect of the environmental or pollutant factors on the degradation is less pronounced (23.63%); (2) the classification accuracy of the trained model on the training set is 91% (91/100), and on the testing set is 72.13% (44/61); and lastly, (3) it is inferred from the error analysis of the classification results that the performance change of polycarbonate in Qionghai and Wuhan is characterized by an initial reduction followed by a slight improvement. Lastly, we show that the proposed method performs well, especially in the case of areas with incomplete data available.

1. Introduction

Polycarbonate (PC) is a widely used engineering plastic owing to its excellent mechanical properties and low specific gravity. However, the deterioration of PC materials is inevitable and largely depends on the ambient factors in their application, e.g., solar radiation, temperature, water exposure, and atmospheric pollution [1–4]. Besides its physical and chemical properties, the weathering of PC is the result of the combined action of various environmental factors. The interplay between these environmental factors is so intricate that the service-life prediction of PC products in many different environments is extremely challenging. Hulme and Cooper [5] concluded that the difficulties of the predictions of the service-life of a polymer are as follows: (1) polymers are time-, temperature-, environment-, and stress-dependent; (2) the limit(s) of various properties of the polymers at which they fail is often unknown; (3) the service conditions generally vary and often include fault situations; and (4) for complex applications, it is impossible to fully replicate the service condition in accelerated tests. However, when information about the time, temperature, environmental factors, and mechanical properties of the polymer can be collected and analyzed on a large scale, the above-listed challenges can be mitigated.

Artificial intelligence (AI) has been rapidly developing in the recent years. Consequently, software supporting various AI tasks is continuously being made available. However, only a few people use this software to apply the latest machine learning methods to study material weathering mechanisms. The limited information available about new algorithms and the lack of skills required to use them may be one of the reasons. Nevertheless, with additional basic tools, the state-of-the-art AI algorithms can be deployed through Python, which is a high-level language suitable for scientific and engineering applications. The use of Python enables the rapid and flexible development of AI applications, which can be further enhanced with additional extensions [6]. In addition, the

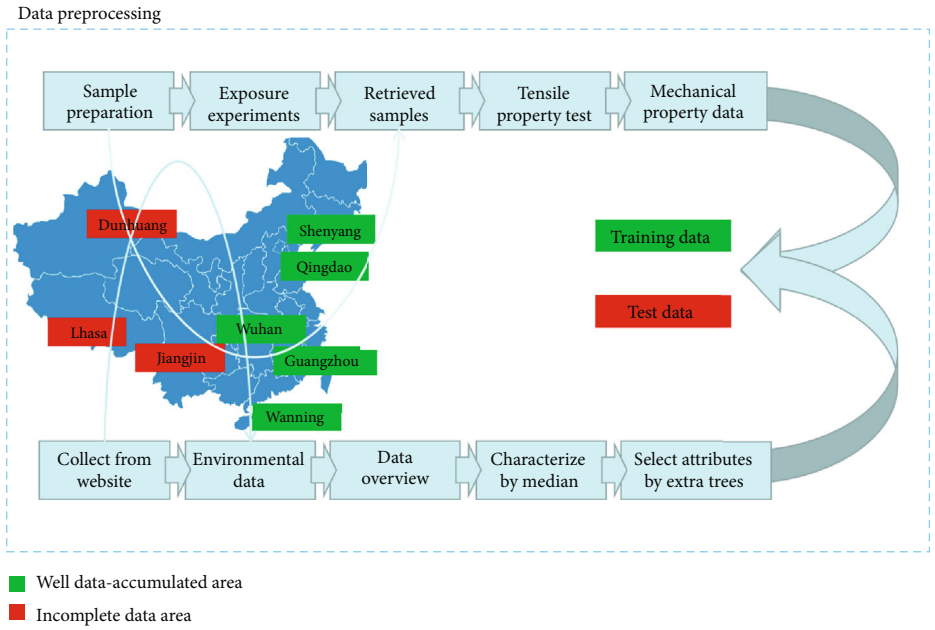

Figure 1: Schematic overview of the preprocessing pipeline.

Python programming language established itself as one of the most popular languages for scientific computing [7]. Thanks to its high-level interactive nature and its maturing ecosystem of scientific libraries, Python is an appealing choice for algorithmic development and exploratory data analysis [8, 9]. Python is easy to learn and convenient to apply. Hence, in the present work, two common machine learning methods (extra-trees and multilayer perceptron networks) have been used for the analysis of various environmental factors and mechanical properties of PC materials.

In fact, the use of Python—or other tools—to integrate machine learning methods for scientific applications has been gaining attention. Ong et al. [10] developed the Python Materials Genomics (pymatgen) library, a robust, open-source Python library for materials analysis. Nevertheless, few researchers have applied Python as the main means of studying weathering mechanisms. Similarly, with the help of machine learning, we can find hidden connections in large datasets.

The study of weathering mechanisms remains a significant and valuable research subject. Many researchers have studied weathering mechanisms through various laboratory-based methods, both at the macro and micro levels [11–14]. However, for enhanced material protection in practical applications, it is necessary to study the weathering mechanisms of materials outdoors. Liu et al. [15] developed an outdoor weathering-life prediction system for PC based on Artificial Neural Network (ANN).

In this work, a three-year-long PC natural weathering test was conducted at eight exposure stations in China. After a tensile-strength and elongation at break factor analysis of the weathered PC, we counted the frequency distribution of the values of all environmental factors in a large-scale data analysis process to identify the most influential factors. Consequently, we separated the most important factors by an extra-trees algorithm to reduce the disturbance introduced of unrelated factors. Lastly, a multilayer perceptron neural network was constructed based on the relationships observed among the characteristic environmental parameters, tensile property variation parameters, and service lifetime. We introduced a guiding action into the model to study weathering mechanisms. There are two main advantages of this method: on the one hand, it maximizes the information extracted from the collected data even if the original data does not have a uniform scale and is incomplete. On the other hand, the applied method can identify macroscopic laws based on the large-scale data analysis.

2. Materials and Methods

2.1. Materials and Sample Preparation. Raw PC materials (K1300, Teijin Limited) were purchased. Standard dumbbell tensile samples (150 mm gauge length, 4 mm × 10 mm cross-section) of pure PC were injection molded on a UA120A injection-molding machine (Yizumi, China). The injecting temperature, mold temperature, injection pressure, packing pressure, and pressure-holding time were 190°C, 40°C, 700 bar, 150 bar, and 10 s, respectively.

2.2. Outdoor Weathering Experiments. According to the ISO 877 standard, the exposure tests were conducted at eight natural exposure stations in China. The stations were exposed to different climate types. The eight stations are located in Wuhan (WH, subtropical zone, humid urban climate type), Lhasa (LS, warm temperate zone, plateau rural climate type), Wanning (WN, torrid zone, marine climate type), Dunhuang (DH, warm temperate zone, dry and hot desert climate type), Shenyang (SY, warm temperate zone, humid urban climate type), Jiangjin (JJ, subtropical zone, suburban acid rain climate type), Guangzhou (GZ, subtropical zone, humid urban climate type), and Qingdao (QD, temperate zone, marine climate type) (Figure 1). The mean monthly values of the main

Table 1: Monthly data of environmental factors, including the temperature (monthly maximum, minimum, and mean), relative humidity (monthly maximum, minimum, and mean), atmospheric pressure (monthly maximum, minimum, and mean), average wind velocity, infrared radiation, ultraviolet radiation, total solar radiation, sunshine duration, precipitation, and rainfall duration in Wuhan, China, 2013.

		Data of meteorological factors of Wuhan experimental station in 2013													
Monthly	Factor	January	February	March	April	May	June	July	August	September	October	November	December	Total	Mean
Temperature (°C)	Monthly mean	5.50	8.90	12.20	17.50	23.00	27.70	30.50	30.30	25.70	20.00	10.70	7.40	219.40	33.75
	Monthly high	10.40	12.70	17.50	23.30	28.70	33.60	25.60	36.20	31.50	26.90	16.90	12.80	286.10	44.02
	Monthly low	1.40	5.60	7.60	13.00	18.40	23.10	26.70	26.00	21.60	15.20	6.00	2.90	167.50	25.12
Relative humidity (%)	Monthly mean	63.30	72.70	67.30	65.80	68.30	68.70	70.80	68.50	70.90	66.30	66.20	64.10	612.90	125.06
	Monthly high	76.40	80.60	80.00	79.00	81.30	81.80	80.80	79.80	80.90	80.60	78.90	76.40	956.50	147.15
	Monthly low	48.80	62.40	53.30	49.30	52.50	52.90	57.30	53.50	56.60	47.40	49.50	49.30	633.00	97.38
Air pressure (hPa)	Monthly mean	1026.90	1016.60	1013.30	1005.90	995.60	986.00	982.10	985.00	994.00	1005.20	1019.60	1021.40	12051.60	1854.09
	Monthly high	1032.00	1021.70	1019.50	1011.20	1000.00	990.90	985.80	989.3	998.10	1009.80	1024.90	1026.70	12109.20	1862.95
	Monthly low	1020.50	1010.80	1005.70	998.50	989.20	979.10	976.10	977.60	987.40	997.10	1011.30	1014.10	11967.40	1841.14
Average wind velocity (m/s)		12.80	13.70	18.10	14.90	14.10	14.30	25.40	22.40	8.20	4.70	9.45	12.40	168.45	25.92
Solar radiation (MJ/m)	Infrared radiation	136.78	84.86	156.59	165.12	162.91	168.91	178.23	208.479	163.93	185	160.09	156.73	1928.09	296.63
	Ultraviolet radiation	17.27	11.78	23.11	28.63	27.39	31.45	34.36	33.76	24.70	23.31	18.75	17.65	292.16	44.95
	Total radiation	233.069	153..50	288.39	320.93	323.42	351.88	370.50	421.33	320.47	336.12	279.58	267.34	3667.13	564.17
Sunshine duration (h/month)		316.40	334.70	339.60	425.10	400.60	409.20	432.80	454.70	434.80	468.00	406.30	378.70	4800.90	738.60
Precipitation (mm/month)		27.00	76.80	85.90	146.90	136.60	251.50	218.00	83.50	62.10	57.10	45.80	32.90	1224.10	188.32
Rainfall duration (h/month)		8.80	21.10	22.50	29.20	29.30	24.30	22.00	9.90	12.90	13.40	10.60	11.40	215.40	33.14

Table 2: Monthly data of the pollutant factors, such as the concentration of sulfur dioxide and hydrogen chloride determined using the instantaneous method; the concentration of nitrogen dioxide, hydrogen sulfide, ammonia, sea salts, and sulfation rate determined using the continuous method; the pH, concentration of sulfate ions and chloride ions in the rain; and the water-soluble and non-water-soluble dust fall quantity in Wuhan, China, 2007.

Month	Data of pollutant factors of Wuhan experimental station in 2007											
	Instantaneous methods (mg/m)		Continuous methods (mg/100 cm^2/day)					Rainfall (mg/m^3)			Dustfall (g/m^2/month)	
	Sulfur dioxide	Hydrogen chloride	Nitrogen dioxide	Hydrogen sulfide	Sulfation rate	Ammonia	Sea salts	pH	Sulfate ion	Chloridion	Water-soluble	Non-water-soluble
January	0.0795	0.0138	0.144	0.1153	0.0434	0.0057	0.0083	6.71	31278.05	1304.94	4.2258	3.3492
February	0.0204	0.0281	0.1443	0.1048	0.7399	0.0178	0.0074	6.51	15616.37	2207.12	4.6604	4.8903
March	0.0496	0.017	0.1767	0.0073	0.6159	0.0468	0.0071	5.6	16799.74	1354.72	6.521	3.0319
April	0.0149	0.0362	0.1317	0.1435	0.3321	0.07	0.0137	6.612	7823.14	—	3.5711	2.5987
May	0.024	0.0164	0.1327	0.1402	0.2875	0.128	0.0158	3.528	16485.26	2243.1	12.6885	12.6885
June	0.0256	—	0.1279	0.0806	0.1282	0.0751	0.013	6.409	7665.72	332.46	2.5339	4.5327
July	0.1324	0.0161	0.0908	0.0223	0.7752	0.1106	0.015	6.425	3999.09	1060.42	4.9648	2.08
August	0.046	0.0269	0.1051	0.0334	0.42	0.0972	0.0115	6.187	8149.1	962.9	3.9206	6.5422
September	0.0483	0.0166	0.1227	0.0335	0	0.0432	0.0078	6.051	12744.48	1098.18	1.7866	3.0268
October	0.089	0.0153	0.1327	0.0721	0.0846	0.042	0.0117	6.704	18227.64	1546.98	2.1519	4.5428
November	—	—	0.0791	0.0407	0.1981	0.0111	0.0079	5.967	23816.26	607.3	2.6715	3.5607
December	0.1281	0.0082	0.0792	0.1986	0.4736	0.0139	0.0064	—	—	839.54	1.1105	2.0397
Total	0.6578	0.1946	1.4669	0.9923	4.0985	0.6614	0.1256	69.7030	162604.8500	13557.6600	50.8066	52.8835
Mean	0.0598	0.0195	0.1222	0.0827	0.3415	0.0551	0.0105	6.3366	14782.2591	1232.5145	4.2339	4.4070

Input: Training Set $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$;
Attribute Set $A = \{a_1, a_2, \cdots, a_d\}$.
Process: Function $TreeGenerate(D, A)$
01: generate the node
02: **if** all samples in D belong to the same class C **then**
03: sign the node as a leaf node of class; **return**
04: **end if**
05: **if** $A = \emptyset$ **or** samples in D have the same value of A **then**
06: sign the node as a leaf node of the class which the most samples in D belong to; **return**
07: **end if**
08: select the optimal attribute;
09: **for** every value a_v^* of a^* **do**
10: generate a branch for the node;
make D_v be a subaggregate of samples in D with the value a_v^* of;
11: **if** $D_v = \emptyset$ **then**
12: sign the branch node as a leaf node of the class which the most samples in D belong to; **return**
13: **else**
14: sign $TreeGenerate(D_v, A, \{a^*\})$ as a branch node
15: **end if**
16: **end for**
Output: a decision tree with the root node.

Algorithm 1

Split_a_node(S)
Input: the local learning subset S corresponding to the node we want to split
Output: a split $[a < a_c]$ **or** nothing
–**If Stop_split**(S) is TRUE **then return** nothing.
–**Otherwise** select K attributes $\{a_1, \cdots, a_k\}$ among all non-constant (in S) candidate attributes;
–Draw K splits $\{s_1, \cdots, s_K\}$, where S_i = **Pick_a_random_split**(S, a_i), $\forall i = 1, \cdots, K$;
–**Return** a split $s*$ such that Score($s*, S$) = $\max_{i=1,\cdots,K}$Score(s_i, S).

Algorithm 2

Pick_a_random_split(S, a)
Inputs: a subset S and an attribute
Output: a split
–Let $a_{\max}^S$ and $a_{\min}^S$ denote the maximal and minimal value of a in S;
–Draw a random cut-point a_C uniformly in $[a_{\min}^S, a_{\max}^S]$;
–**Return** the split $[a < a_C]$.
Stop_split(S)
Input: a subset S
Output: a boolean
–**If** $|S| < n_{\min}$, **then return** TRUE;
–**If** all attributes are constant in S, **then return** TRUE;
–**If** the output is constant in S, **then return** TRUE;
–**Otherwise, return** FALSE.

Algorithm 3

environmental factors—such as monthly mean, high, and low temperature (T); monthly mean, high, and low relative humidity (RH); rainfall duration (RD); precipitation (P); sunshine duration (S); total solar radiation (G); infrared radiation (IR); ultraviolet radiation (UV); sulfur dioxide (SD); hydrogen chloride (HC); nitrogen dioxide (ND); hydrogen sulfide (HS); sulfate ion rate (SR); ammonia (A); sea salts (SS); pH; sulfate ion (SI); chloride (Cl); and dust fall (DF)—of the exposure sites are supplied in the Supplementary materials (available here). The data of the environmental parameters were collected and calculated from the National Material Environmental Corrosion Platform. During the three-year duration of the

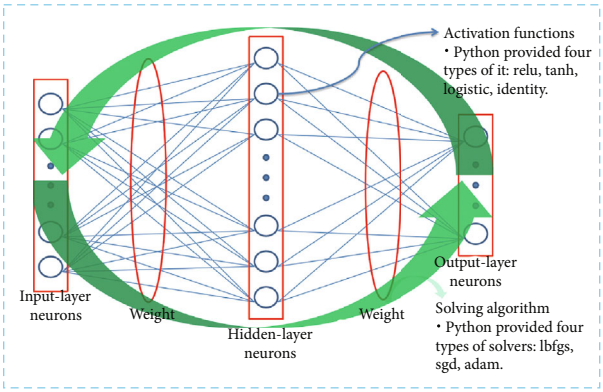

Figure 2: Basic principles and structure of the multilayer perceptron networks implemented in Python.

outdoor weathering, the PC samples were fixed at both ends on aluminum alloy frames, tilted at 45° from the horizontal position, and directly exposed to the south without any backing.

2.3. Characterization Methods. Dumb-bell specimens were used for tensile tests according to the ISO 527.2 standard, using a universal material machine (CMT 6503, MTS Systems Corporation), at a stretching rate of 20 mm/min.

2.4. Data Source. All environmental-factor data were collected from the website http://data.ecorr.org, which was built for public research. The original data was recorded in excel tables such as the ones shown in Tables 1 and 2. We focused on five areas whose environmental factors were fully documented. We will refer to these areas as well data-accumulated areas (Guangzhou, Qingdao, Shenyang, Wanning, and Wuhan). The remaining three areas will be referred to as incomplete data areas (Dunhuang, Jiangjin, and Lhasa).

The data used in the presented experiments were recorded during various intervals: between 2005 and 2012 and in 2014 in Qingdao and Wanning, between 2006 and 2014 in Shenyang and Wuhan, between 2005 and 2014 in Guangzhou, and between 2012 and 2014 in Dunhuang, Jiangjin, and Lhasa. The main part of the experiment relied on the data collection and preprocessing, as shown in Figure 1.

2.5. Data Analysis Implemented in Python

2.5.1. Attribute Selection with Extra-Trees. To reduce the error of the subsequent mathematical model, it is necessary to identify the most important factors affecting the outdoor weathering process of PC. Hence, we carried out attribute selection with the extra-trees algorithm integrated in the *sklearn* library.

The extra-trees algorithm was developed from random decision trees, which is a classic machine learning method. The traditional decision trees divided all objects into different branches according to whether their characteristics fit the filtering conditions of the individual branches. The basic process of traditional decision trees is shown in Algorithm 1.

Geurts et al. [16] developed the extra-trees algorithm by adding the following process, which extremely increases the randomness of decision trees.

For regression problems [16], we have

$$\text{Score}(s_i, S) = \frac{\text{var}\{y|S\} - (S_i/|S|)\,\text{var}\{y|S_l\} - (S_i/|S|)\,\text{var}\{y|S_r\}}{\text{var}\{y|S\}}, \quad (1)$$

where var $\{y|S\}$ refers to the variance of the output y in sample set S, r and l refer to the right and left branch of the node, respectively.

Moreover, Pierre Geurts proved that the extra-trees learning algorithm could provide near-optimal accuracy and good computational complexity, especially on classification problems.

The criteria for selecting optimal decisive attributes is the key to successful sample classification. Similarly, if applying a given attribute as a decisive attribute can significantly improve the classification accuracy, that attribute is important for describing the weathering mechanisms of the samples. Therefore, multiple methods are available for selecting the optimal and decisive attributes.

Geurts et al. [17] have described the variable importance in forests of randomized trees. They demonstrated that Mean Decrease Impurity (MDI) importance computed by totally randomized trees and extra-trees exhibit desirable properties for assessing the relevance of a variable: it is equal to zero if and only if the variable is irrelevant and it depends only on the relevant variables.

We used this method to find the important factors which can contribute more significantly to the weathering process of PC materials than other factors.

2.5.2. Weathering-Life Prediction Model with Multilayer Perceptron Networks. With the development of artificial intelligence in the recent years, ANNs gained widespread popularity for data processing in various industries. There is an increasing number of convenient computer software packages that facilitate the use, implementation, and application of ANNs. For example, only by importing several parameters can a corresponding model be implemented in Python, which was used in the current work.

The structure and basic principles of multilayer perceptron networks—as one of the first ANN models—have already been described by many researchers [18]. Hence, the multilayer perceptron network shown in Figure 2 was implemented in Python.

Besides its basic use, the *sklearn* library provides alternative activation functions and solving algorithms (the solvers for weight optimization) which work in various situations. For example, "lbfgs" is an optimizer from the family of quasi-Newton methods; "sgd" refers to stochastic gradient descent; "adam" refers to the stochastic gradient-based optimizer proposed by Kingma and Ba [19]. In addition, it is noted in the help document of Python that the default solver "adam" works well on relatively large datasets (with thousands of training samples or more) in terms of both training time and validation score. For small datasets, however, "lfbgs" can converge faster and perform better. Hence, we chose the "lfbgs" solver in this work.

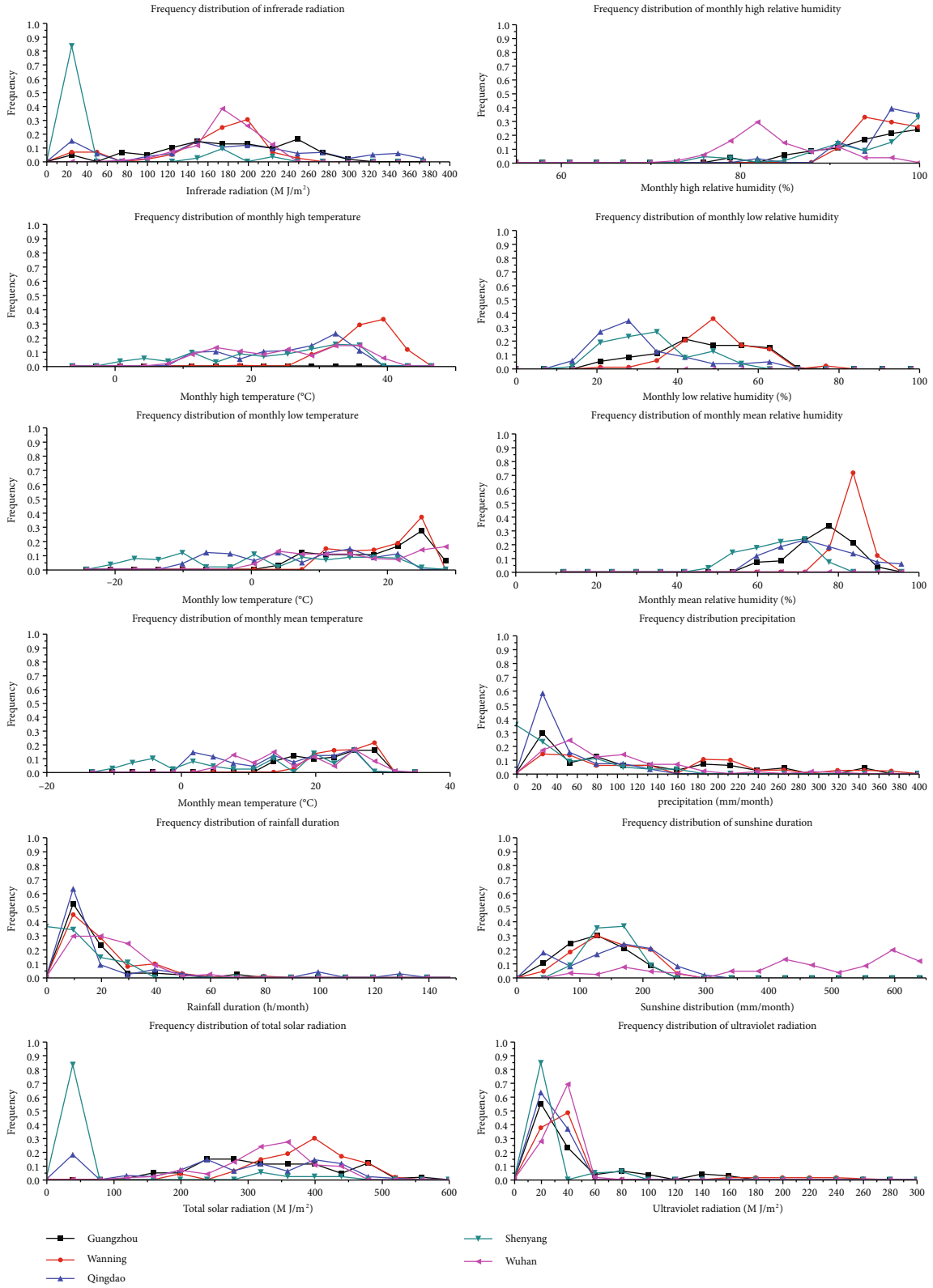

Figure 3: Frequency distribution graphs of climatic factors.

3. Results

3.1. Data Overview and Characterization. As Tables 1 and 2 show, there are 28 recorded factors in total, including 16 climatic factors and 12 pollutant factors. According to the quality of the existing data and weathering mechanism models of polymer materials, 7 factors (air pressure, wind velocity and direction, solar radiation angel, the pH, concentration of

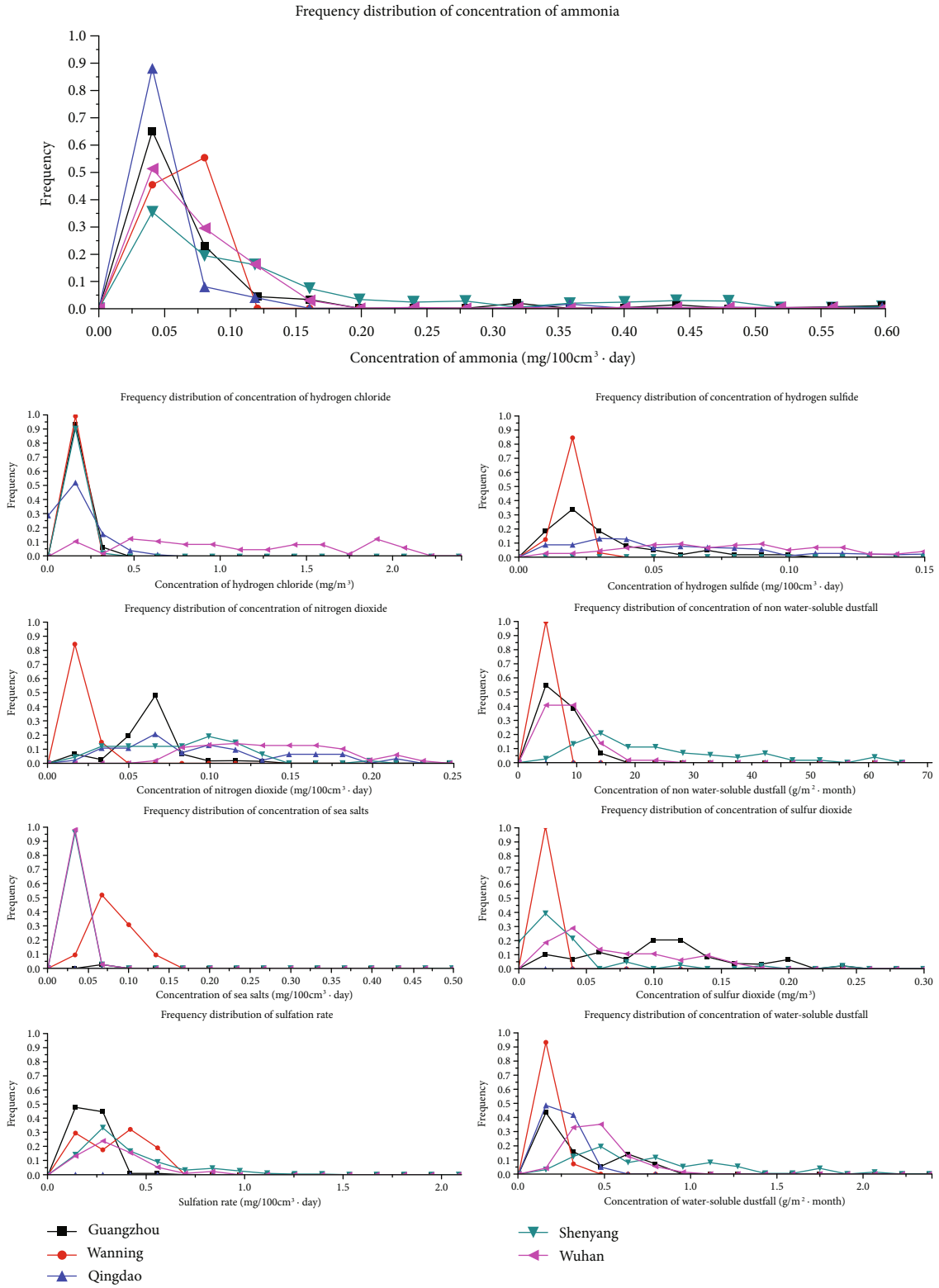

Figure 4: Frequency distribution graphs of pollutant factors.

sulfate ions and chloride ions in the rain) were not used in this work owing to the high level of noise in the data associated with these factors.

Some data were not recorded due to reasons beyond the experimental control. Consequently, more accurate conclusions could be drawn from a complete dataset. Therefore, a method that was robust against small numbers of missing values is needed.

Firstly, frequency distribution graphs were constructed (Figures 3 and 4) for the whole dataset, i.e., for each factor,

Table 3: Median values of the pollutant factors in the well data-accumulated regions.

Factors	Regions				
	Guangzhou	Wanning	Qingdao	Shenyang	Wuhan
Sulfur dioxide (mg/m^3)	0.096	0.002	0	0.0082	0.0421
Hydrogen chloride (mg/m^3)	0.0946	0.0638	0.0399	0.0203	0.9288
Experimental period (month)	0.0563	0.0101	0.08	0.0769	0.1314
Nitrogen dioxide ($mg/100\,cm^2{\cdot}day$)	0.0166	0.0131	0.0521	0.0655	0.079
Hydrogen sulfide ($mg/100\,cm^2{\cdot}day$)	0.1456	0.2867	0.2846	0.5118	0.279
Sulfation rate ($mg/100\,cm^2{\cdot}day$)	0.0348	0.0431	0.0176	0.0664	0.0394
Ammonia ($mg/100\,cm^2{\cdot}day$)	0.0098	0.0584	0.1645	0.0115	0.0089
Sea salts ($mg/100\,cm^2{\cdot}day$)	2.083	0.7638	1.6562	6.6289	3.719
Water-soluble dust (g/m^2)	4.2203	3.063	2.9046	18.6164	5.1195

Table 4: Median values of the climatic factors in the well data-accumulated regions.

Factors	Regions				
	Guangzhou	Wanning	Qingdao	Shenyang	Wuhan
Monthly mean temperature (°C)	24.08	26.2	13.75	12.2	18.8
Monthly high temperature (°C)	32.7	35.6	25.5	23	24.5
Monthly low temperature (°C)	18.42	19.8	4.95	-0.8	14.2
Monthly mean humidity (%)	74.75	81.18	71.5	63	68.5
Monthly high humidity (%)	93.95	95	97	94	81.3
Monthly low humidity (%)	44	46.25	24	29	53.4
Infrared radiation (MJ/mm^2)	172.87	165.47	171.9	6.68	168.34
Ultraviolet radiation (MJ/mm^2)	17.27	22.58	15.45	0.89	24.23
Total solar radiation (MJ/mm^2)	312.1	367.1	253.3	12.09	314.67
Sunshine duration (h/month)	110.3	123.05	138.1	130.45	442.4
Precipitation (mm/month)	77.6	140.6	18.6	15.8	74.4
Rainfall duration (h/month)	8.27	12	5	3.5	15.95

so that the data could be intuitively presented. A frequency distribution chart of all environmental factors was drawn to observe the distribution of different factors in each region, and to determine whether median values could be used to suitably characterize and simulate the actual environment. There are similarities to study the impact of individual key factors on performance data. For calculation, after uniformly dividing the maximum and minimum values of each factor into 15 intervals, based on the amount of data falling in different intervals, the frequency of the total data of the factor is determined.

Based on the frequency distribution graphs, the data about most factors were concentrated in a specific interval that depended on the region. The temperature and relative humidity appear relatively as noise as these two climatic factors periodically vary with the seasons. However, there is a distinct difference between the curve fluctuations, suggesting that this noise would not affect the distinction between the environments of distinct regions. Hence, the median values of every factor were selected to characterize them based on the regions, as shown in Tables 3 and 4.

3.2. Importance of Factors. The ExtraTreesClassifier package of the *sklearn* library of Python 3.6 offers a convenient approach to determining the importance of each factor, as shown in Tables 5 and 6. When using the algorithm, we set the environmental factors—including the pollutant factors and the experimental time periods—as the cause and set the elongation of materials at break as the result. Subsequently, the ExtraTreesClassifier operation was repeated 100 times. The mean and the variance of the 100 results were used to determine the importance of each factor with statistical significance.

Tables 5 and 6 suggest that the experimental period is the most important factor as it exhibited an importance of 0.7637 that is one or two orders larger than that of the other factors. Hence, the experimental period contributes most significantly to the tensile property degradation of PC materials (76.37%). This is in line with most laboratory findings. Besides, the remaining approximately 25% importance can be attributed to the environmental factors. These relationships are shown in Figure 5.

Based on existing research, high temperature can affect the mechanical properties of polymer materials. Hence, all factors with importance parameter larger than the monthly maximum temperature were chosen for the next step. Table 7 shows the data of the chosen factors in regions with partially missing data. Eventually, the data was combined

Table 5: Mean and variance of the importance of the pollutant factors after 100 repetitions of the calculations.

Factors	Mean of the 100 results	Variance of the 100 results
Sulfur dioxide (mg/m^3)	0.00969447	3.55*E*-05
Hydrogen chloride (mg/m^3)	0.01319386	5.18*E*-05
Experimental period (month)	0.76373876	6.70*E*-03
Nitrogen dioxide (mg/100 cm^2·day)	0.01498199	7.16*E*-05
Hydrogen sulfide (mg/100 cm^2·day)	0.00904402	3.97*E*-05
Sulfation rate (mg/100 cm^2·day)	0.00600798	2.41*E*-05
Ammonia (mg/100 cm^2·day)	0.01206354	4.74*E*-05
Sea salts (mg/100 cm^2·day)	0.02162541	8.38*E*-05
Water-soluble dust (g/m^2)	0.0084063	3.89*E*-05
Non-water-soluble dust (g/m^2)	0.0059862	2.42*E*-05

Table 6: Mean and variance of the importance of the climatic factors after 100 repetitions of the calculations.

Factors	Mean of the 100 results	Variance of the 100 results
Monthly mean temperature (°C)	0.00852239	3.15*E*-05
Monthly high temperature (°C)	0.00914544	3.62*E*-05
Monthly low temperature (°C)	0.00712514	2.77*E*-05
Monthly mean humidity (%)	0.00885633	4.42*E*-05
Monthly high humidity (%)	0.01877014	9.72*E*-05
Monthly low humidity (%)	0.01216982	4.30*E*-05
Infrared radiation (MJ/mm^2)	0.00647525	3.20*E*-05
Ultraviolet radiation (MJ/mm^2)	0.00876974	3.95*E*-05
Total solar radiation (MJ/mm^2)	0.00665031	2.21*E*-05
Sunshine duration (h/month)	0.01466707	6.16*E*-05
Precipitation (mm/month)	0.01137561	5.17*E*-05
Rainfall duration (h/month)	0.01273025	4.96*E*-05

with the mechanical properties of the materials tested in those regions.

3.3. Training and Testing. For the training dataset, to predict the outdoor service-life of PC materials, we set the experimental period as the output and 14 factors (the 11 environmental factors in Table 7 and 3 mechanical properties: tensile strength, yield strength, and elongation at break) as the input. In order to obtain the best performance, every combination (especially the hidden-layer structure and the activation function) within reasonable limits has been explored. Figure 6 shows the results. Figure 6(a) is the identity activation function; the best hidden-layer structure is 27-38, and the achieved accuracy is 62.30%. Figure 6(b) is the logistic activation function; the best hidden-layer structure is 12-23, and the achieved accuracy is 70.49%. Figure 6(c) is the relu activation function; the best hidden-layer structure is 15-16, and the achieved accuracy is 72.13%. Figure 6(d) is the tanh activation function; the best hidden-layer structure is 10-31, and the achieved accuracy is 70.49%.

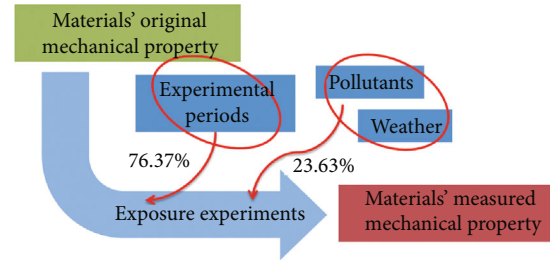

Figure 5: Relationship schema between the various factors influencing the tensile property degradation.

Overfitting is a common problem in machine learning models. Therefore, the accuracy achieved on the training data set cannot be a reference standard of the models' performance. Figure 6 shows the accuracy of four types of activation functions for several hidden-layer structures and shows that the accuracy can reliably measure models' performance. Furthermore, the figure shows that the optimal combination is the "relu" activation function.

The neural network with two hidden layers with 15 and 16 neurons, respectively, combined with the "relu" activation function provided the best classification accuracy: 91% on the training set and 72.13% on the testing set.

4. Discussions

The model trained by 100 input features can recognize the training data with a 91% accuracy (91/100) and the test data with a 72.13% accuracy (44/61). Moreover, additional information can be extracted from the specifics of the recognitions. Tables 8 and 9 show the classification accuracy for various well data-accumulated areas and incomplete data areas, respectively. Tables 10 and 11 show the classification accuracy of various experimental periods in the well data-accumulated areas and the incomplete data areas, respectively.

In terms of the recognition of the training data, the observations in Qingdao and Shenyang (40 observations) were all accurately classified, whilst a single false recognition (8/9) was observed for Wanning and Wuhan (Table 8). The single false recognition resulted from a confusion between 12 months and 36 months whilst no recognition was confused between 12 months and 24 months, including the incomplete data regions (Table 11). Therefore, it can be inferred that the outdoors mechanical degradation of polycarbonate is characterized by an initial deterioration followed by a slight improvement (as shown in Figure 7) assuming that the systematic error generated during data collection for the three different experimental periods is the same.

There is a significant deterioration of the mechanical properties between the samples exposed to the environment for 12 months and 24 months. Hence, no false recognition was observed among them. Almost all misclassifications (25/26) are related to the 36-month samples. Hence, it is probable that the mechanical properties of the samples

Table 7: Data of the chosen factors from the regions with incomplete data.

Factors	Dunhuang	Median Jiangjin	Lhasa
Sulfur dioxide (mg/m^3)	0.00765	0.1786	0.00155
Hydrogen chloride (mg/m^3)	0.0817	0.0018	0
Nitrogen dioxide (mg/100 cm^2·day)	0.05945	0.0288	0.004284
Ammonia (mg/100 cm^2·day)	0.0062	0.05685	0.01545
Sea salts (mg/100 cm^2·day)	0.4321	0.0029	0.002835
Monthly high temperature (°C)	25.85	28.05	20.85
Monthly high humidity (%)	72	100	93
Monthly low humidity (%)	11.5	35	5
Sunshine duration (h/month)	224.25	68.5	184.19
Precipitation (mm/month)	0.05	63.85	25.95
Rainfall duration (h/month)	0	36.55	4.095

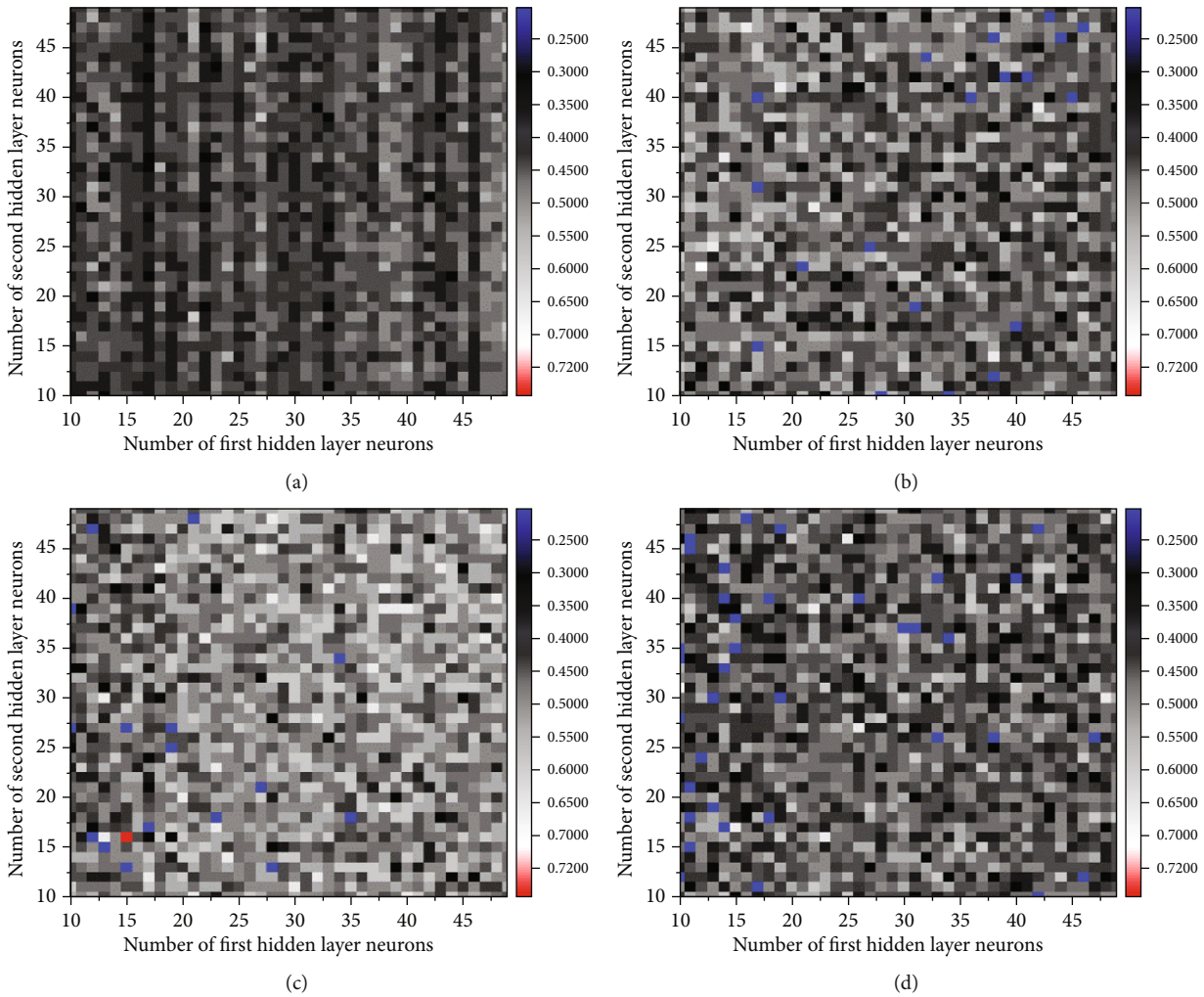

Figure 6: Performance of the explored neural network structures with the number of neurons in the two hidden layers varying between 10 and 50.

exposed for 36 months fall between those of the samples exposed for 12 months and 24 months. Therefore, it is difficult to accurately classify the samples exposed for 36 months. The frequency of the misclassification of the samples exposed for 36 months as samples exposed for 12 months or as samples exposed for 24 months depended on whether the mechanical properties of the samples exposed for 36 months were closer to those of the samples exposed for 12 months or

Table 8: Classification accuracy of the experimental region in various well data-accumulated areas.

	Guangzhou	Qingdao	Shenyang	Wanning	Wuhan	Total
True	18/19	20/20	20/20	18/21	15/20	91/100
False	1/19	0/20	0/20	3/21	5/20	9/100

Table 9: Classification accuracy of the experimental regions for various incomplete data areas.

	Dunhuang	Jiangjin	Lhasa	Total
True	12/20	17/20	15/21	44/61
False	8/20	3/20	6/21	17/61

Table 10: Classification accuracy of the experimental periods in the well data-accumulated areas.

	×12	×24	×36
$\sqrt{12}$	—	0	2
$\sqrt{24}$	0	—	0
$\sqrt{36}$	6	1	—

Table 11: Classification accuracy of various experimental periods in incomplete data areas.

	×12	×24	×36
$\sqrt{12}$	\	0	0
$\sqrt{24}$	1	\	3
$\sqrt{36}$	5	8	\

samples exposed for 24 months. The statistics obtained from every sample shown in Figure 7 support these findings.

Considering the results shown in Figure 8, the black line is more likely to coincide with the blue line than the red one. Moreover, the black line intersects the blue line at 73%. The crossover point at 73% also divides the three lines into two parts: on the right side of the crossover point, all three lines exhibit the same trend, reaching their maxima at 100%; on the left side of the crossover point, the red line differs from the other two by exhibiting a semi-circle-shaped peak. This suggests that the samples exposed for three years (three-year samples) have a similar breaking elongation value distribution as the one-year samples. However, the breaking elongation values of the three-year samples will partially shift to the left as a result of the weathering and the degradation of the mechanical properties of the materials. Therefore, a higher proportion of the three-year samples reached an elongation before break above 73% and a lower proportion below 73%. Similarly, the two-year samples have the same value distribution and a significant and concentrated left shift, especially in the part below 73%. Jiang et al. [20] researched the weathering mechanism of bisphenol A polycarbonate. Their results support our findings. Jiang et al. showed that this phenomenon is a weathering-induced ductile-brittle-ductile transition which partially results from the competition between oxidation-induced chain-scission and chain

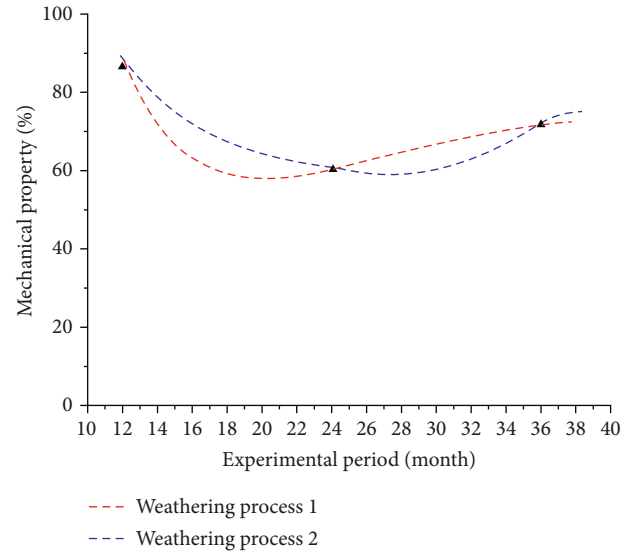

Figure 7: Inferred outdoor mechanical property degradation of polycarbonate.

crosslinking. Hence, it is inferred that our polycarbonate samples also exhibit the same weathering mechanism.

Table 9 shows that the classification accuracy of the experimental regions in incomplete data areas is less than that of the well data-accumulated areas by 18.87%. There are three possible reasons:

(1) The median characterization does not perform well because of the smaller amount of data available in the incomplete data areas. Errors tend to be more pronounced with smaller amounts of data

(2) Statistically, the data range of some factors in the incomplete data areas is far beyond that in the well data-accumulated areas. Hence, the difficulty of recognizing the test data is beyond the capacity of the model trained by the limited training data

(3) Fundamentally, different regional environmental characteristics led to essential differences in the property degradation of the polycarbonate between the training data areas and the test data areas

5. Conclusions

It was proven that by using the integrated tools of Python, it is possible to conveniently analyze data with the state-of-the-art mathematical methods. The important climatic factors and pollutant factors affecting the breakage elongation identified by the extra-trees algorithm present high stability and interpretability. In addition, the important parameters

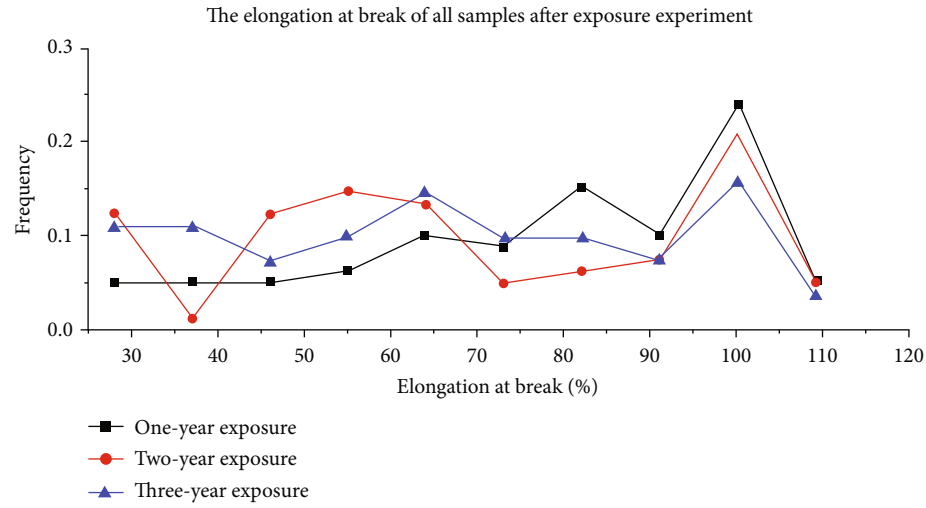

Figure 8: Elongation at break of all samples after the exposure experiments.

guided a more reasonable use of the data in the subsequent process and improved the performance of the multilayer perceptron model. If limited amounts of data are available, looping through all possible combinations with high computing performance is a reliable way to find the optimal hyperparameters of a machine learning model. The model obtained through this method could recognize the experimental periods with relatively high accuracy. This provided an important reference value for the study of weathering processes and appropriate protection measures for polycarbonate in atmospheric environments.

According to the error analysis, from the macroscopic point of view, the outdoors mechanical properties of polycarbonate would deteriorate first and then rise slightly. This suggests that the outdoor weathering process of polycarbonate is a ductile-brittle-ductile transition.

It is feasible to predict the weathering periods of samples in incomplete data areas with samples in the well data-accumulated areas, although with modest errors. Moreover, it is more accurate to predict the service life of certain samples from the data obtained from well data-accumulated areas.

Data Availability

The raw data of materials' mechanical property required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study. The environmental data can be found in China Geteway to Corrosijon and Protection (http://data.ecorr.org/) and a part of them is in the attachment.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

We are grateful for the financial support of the National Natural Science Foundation of China (No. 51133009). We also thank the Mingshu Yang team of the Institute of Chemistry of the Chinese Academy of Sciences for their assistance in sample preparation.

Supplementary Materials

In the file data.xlsx, the sheet name place-1 means the data of the meteorological factors of this place. The sheet name place-2 means the data of the pollutant factors of this place. In all sheets of the file, Guangzhou, Qingdao, Shenyang, and Wuhan are referred to as well data-accumulated areas. The remaining three areas will be referred to as incomplete-data areas (Dunhuang, Jiangjin, and Lhasa). The data used in the presented experiments were recorded during various intervals: between 2005 and 2012 and in 2014 in Qingdao; between 2006 and 2014 in Shenyang and Wuhan; between 2005 and 2014 in Guangzhou; and between 2012 and 2014 in Dunhuang, Jiangjin, and Lhasa. All tables in the file contained monthly data of environmental factors, including the temperature (monthly maximum, minimum, and mean), relative humidity (monthly maximum, minimum, and mean), atmospheric pressure (monthly maximum, minimum, and mean), infrared radiation, ultraviolet radiation, total solar radiation, sunshine duration, precipitation, rainfall duration; monthly data of the pollutant factors, such as the concentration of sulfur dioxide and hydrogen chloride determined using the instantaneous method; the concentration of nitrogen dioxide, hydrogen sulfide, ammonia, sea salts, and sulfation rate determined using the continuous method; the pH, concentration of sulfate ions, and chloride ions in the rain; and the water-soluble and non-water-soluble dust fall quantity in the year. *(Supplementary Materials)*

References

[1] A. Factor and M. L. Chu, "The rôle of oxygen in the photo-ageing of bisphenol-A polycarbonate," *Polymer Degradation and Stability*, vol. 2, no. 3, pp. 203–223, 1980.

[2] R. J. Gardner and J. R. Martin, "Humid aging of plastics: effect of molecular weight on mechanical properties and fracture

morphology of polycarbonate," *Journal of Applied Polymer Science*, vol. 24, no. 5, pp. 1269–1280, 1979.

[3] M. Diepens and P. Gijsman, "Outdoor and accelerated weathering studies of bisphenol A polycarbonate," *Polymer Degradation and Stability*, vol. 96, no. 4, pp. 649–652, 2011.

[4] M. Diepens and P. Gijsman, "Photodegradation of bisphenol A polycarbonate," *Polymer Degradation and Stability*, vol. 92, no. 3, pp. 397–406, 2007.

[5] A. Hulme and J. Cooper, "Life prediction of polymers for industry," *Sealing Technology*, vol. 2012, no. 9, pp. 8–12, 2012.

[6] T. E. Oliphant, "Python for scientific computing," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 10–20, 2007.

[7] A. Swami and R. Jain, "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 10, pp. 2825–2830, 2012.

[8] K. J. Millman and M. Aivazis, "Python for scientists and engineers," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 9–12, 2011.

[9] P. F. Dubois, "Guest Editor's introduction: Python: batteries included," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 7–9, 2007.

[10] S. P. Ong, W. D. Richards, A. Jain et al., "Python Materials Genomics (pymatgen): a robust, open-source python library for materials analysis," *Computational Materials Science*, vol. 68, pp. 314–319, 2013.

[11] J. M. Hutchinson, S. Smith, B. Horne, and G. M. Gourlay, "Physical aging of polycarbonate: enthalpy relaxation, creep response, and yielding behavior," *Macromolecules*, vol. 32, no. 15, pp. 5046–5061, 1999.

[12] J. Bartos, J. Müller, and J. H. Wendorff, "Physical ageing of isotropic and anisotropic polycarbonate," *Polymer*, vol. 31, no. 9, pp. 1678–1684, 1990.

[13] T. Ricco and T. L. Smith, "Rejuvenation and physical ageing of a polycarbonate film subjected to finite tensile strains," *Polymer*, vol. 26, no. 13, pp. 1979–1984, 1985.

[14] A. Blaga and R. S. Yamasaki, "Surface microcracking induced by weathering of polycarbonate sheet," *Journal of Materials Science*, vol. 11, no. 8, pp. 1513–1520, 1976.

[15] H. Liu, M. Zhou, Y. Zhou et al., "Aging life prediction system of polymer outdoors constructed by ANN. 1. Lifetime prediction for polycarbonate," *Polymer Degradation and Stability*, vol. 105, pp. 218–236, 2014.

[16] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine Learning*, vol. 63, no. 1, pp. 3–42, 2006.

[17] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in forests of randomized trees," *Advances in Neural Information Processing Systems*, vol. 2013, pp. 431–439, 2013.

[18] Z. Zhang and K. Friedrich, "Artificial neural networks applied to polymer composites: a review," *Composites Science and Technology*, vol. 63, no. 14, pp. 2029–2044, 2003.

[19] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *Computer Science*, vol. 2014, article 1412.6980, 2014.

[20] L. Jiang, M. Zhou, Y. Ding, Y. Zhou, and Y. Dan, "Aging induced ductile-brittle-ductile transition in bisphenol A polycarbonate," *Journal of Polymer Research*, vol. 25, no. 2, p. 39, 2018.