Supplemental material for

"Impact of Climate Change on Potential Distribution Patterns of Alpine Vegetation in the Hengduan Mountains Region, China", by Yunling He, Qiaoli Xiong, Lan Yu, Wenbo Yan, and Xinxing Qu, published in *Mountain Research and Development* 40(3), 2020. (See https://bioone.org/toc/mred/40/3)

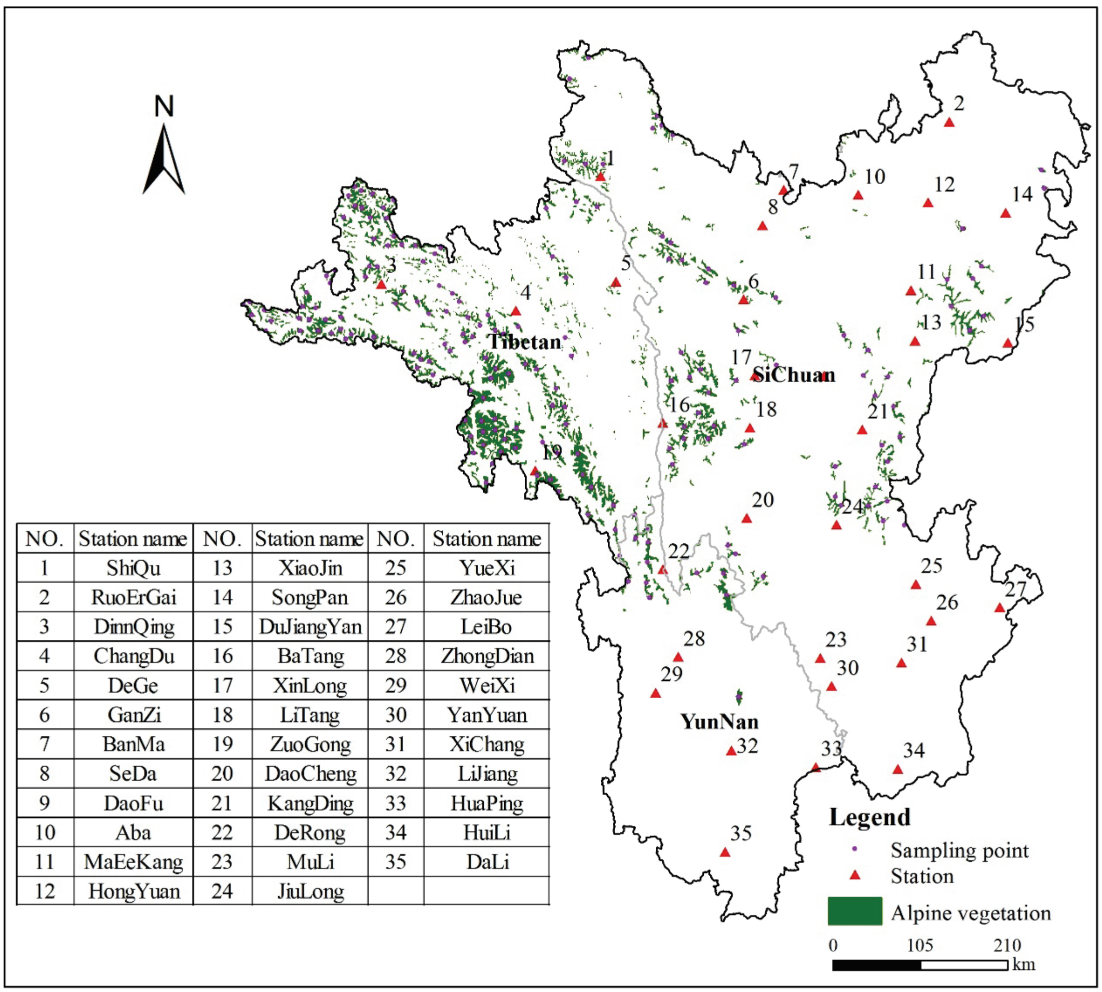
APPENDIX 1 General situation and location details.

Table 1 General characteristics of the selected meteorological stations in the Hengduan Mountains

No.	Station	Longitude (°E)	Latitude (°N)	Elevation (m)
1	56038	98.06	32.59	4764
2	56079	102.58	33.35	3638
3	56116	95.36	31.25	4802
4	56137	97.10	31.09	4608
5	56144	98.35	31.48	4749
6	56146	100.00	31.37	4770
7	56151	100.45	32.56	4084
8	56152	100.20	32.17	4366
9	56167	101.07	30.59	2862
10	56171	101.42	32.54	3178
11	56172	102.14	31.54	2744
12	56173	102.33	32.48	3619
13	56178	102.21	31.00	3310
14	56182	103.34	32.39	3308
15	56188	103.40	31.00	2251
16	56247	99.06	30.00	3106
17	56251	100.19	30.56	4276
18	56257	100.16	30.00	3957
19	56331	97.50	29.40	4347
20	56357	100.18	29.03	4238
21	56374	101.58	30.03	3936
22	56441	99.17	28.43	2132
23	56459	101.16	27.56	3425
24	56462	101.30	29.00	4588
25	56475	102.31	28.39	2360
26	56479	102.51	28.00	2733
27	56485	103.35	28.16	2628
28	56543	99.42	27.50	3051

29	56548	99.17	27.10	1667
30	56565	101.31	27.26	2784
31	56571	102.16	27.54	1863
32	56651	100.13	26.52	2428
33	56664	101.16	26.38	1973
34	56671	102.15	26.39	1709
35	56751	100.11	25.42	2495

Figure 1 Geographical distribution of meteorological stations in the Hengduan Mountains



NO.	Station name	NO.	Station name	NO.	Station name
1	ShiQu	13	XiaoJin	25	YueXi
2	RuoErGai	14	SongPan	26	ZhaoJue
3	DinnQing	15	DuJiangYan	27	LeiBo
4	ChangDu	16	BaTang	28	ZhongDian
5	DeGe	17	XinLong	29	WeiXi
6	GanZi	18	LiTang	30	YanYuan
7	BanMa	19	ZuoGong	31	XiChang
8	SeDa	20	DaoCheng	32	LiJiang
9	DaoFu	21	KangDing	33	HuaPing
10	Aba	22	DeRong	34	HuiLi
11	MaEeKang	23	MuLi	35	DaLi
12	HongYuan	24	JiuLong		

APPENDIX 2 The simulation process of the MaxEnt model.

1. The principles of MaxEnt model

Entropy was proposed by Glausius, a German physicist, as a concept of thermodynamics in 1865. In 1948, C.E.Annon, the father of information theory, first introduced the concept of entropy into information theory, and understood entropy as the probability of occurrence of certain information (Zhang Weibo, 2014). The maximum entropy theory holds that under known conditions, the thing with the maximum entropy is closest to its true state. In the maximum entropy algorithm estimation, the true distribution of species is expressed as the probability distribution P on the set of X sites in the study area (Renner et al., 2013). Therefore, there is a non-negative probability P(x) for each site X, and then model the probability distribution P with the data of species distribution point as the limiting factor. Simulated in species distribution, assuming that randomly selected from a site x1 from the site set x , if there is a species is written down to 1, does not exist, it is 0, y for response variables (existence), the probability distribution of P (x) = P (x | y = 1), namely the species known in the study area distribution situation, get the probability of species x observation in site. According to Bayes' theorem:

$$P(y=1|x) = \frac{P(x|y=1)P(y=1)}{P(x)} \tag{1}$$

In equation (1), P(y=1|x) is the probability that the species is distributed at site x, and P(y=1) is the probability that the species is distributed in the whole region. In practical applications, P(y=1) cannot be obtained only from the observation data of sampling points, so P(y=1|x) cannot be directly estimated and the maximum entropy of P(x) can be estimated.

Therefore, MaxEnt model is a probability model to predict the potential distribution of species based on the principle of ecological niche. In practical application, the MaxEnt model evaluates the habitat suitability of species by using the data of species occurrence points and environmental variables, selects the distribution with the maximum entropy as the optimal distribution from the eligible distribution, and the predicted result is the relative probability of species existence. The output range of the model is 0-1, and the larger the value is, the more suitable the growth of the species is. The wonderful results of using the MaxEnt model can be obtained even in poor conditions, and the modeling results are also very clear and accurate; thus, it has the advantages of simplicity, rapid execution, and good predicting results when compared with other SDM models (Morales et al., 2017).

The MaxEnt model under JAVA running is adopted in this paper, and its operation interface is shown in Figure 1. The full workflow that provided the basis for the analysis is summarized in the Figure 2.

Figure 1 Interface diagram of MaxEnt model operation

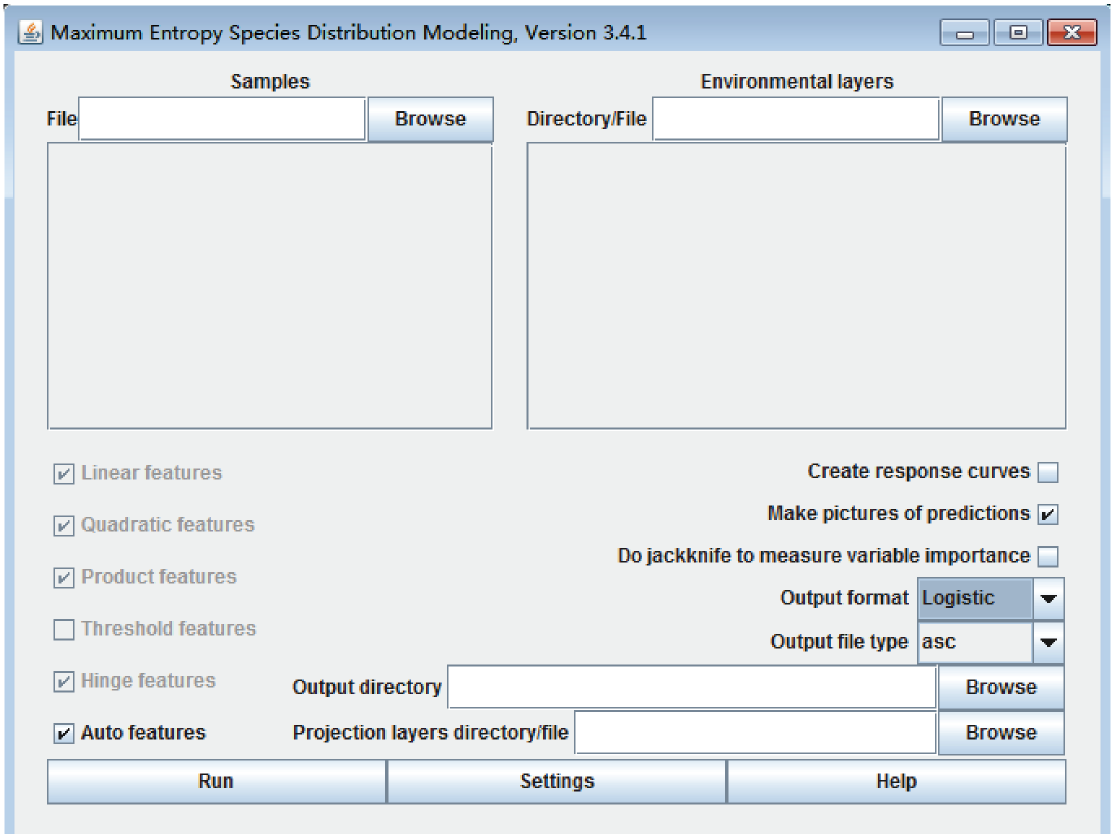
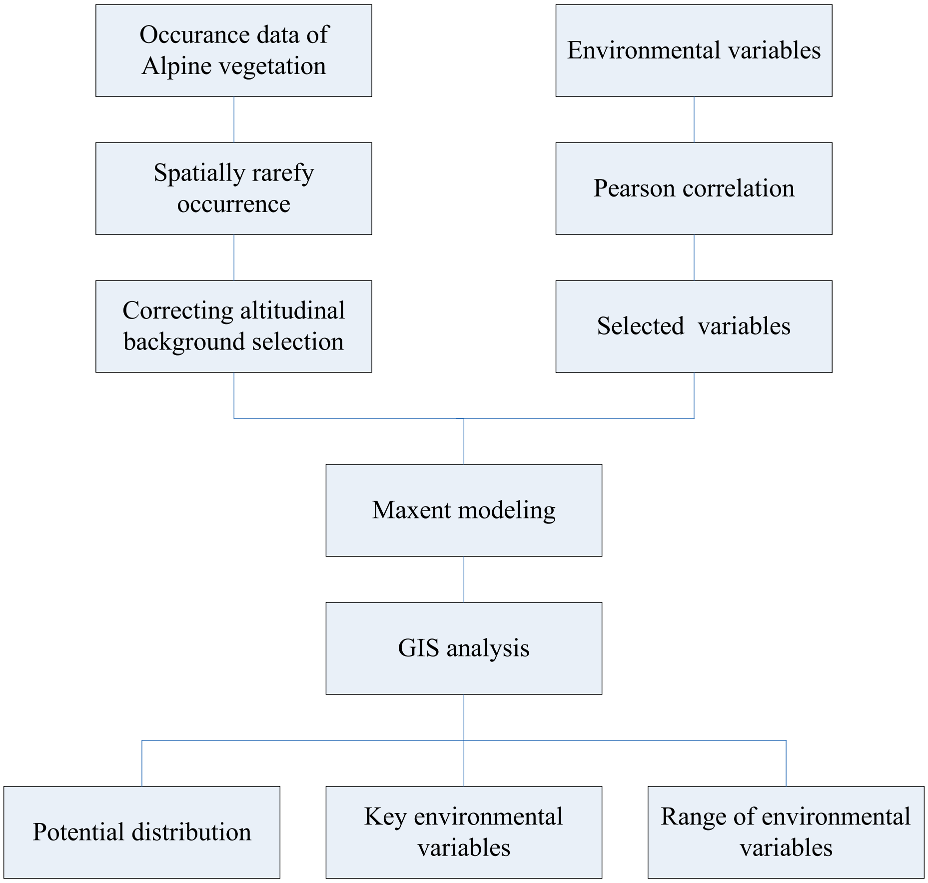


Figure 2 The full workflow for the MaxEnt analysis.



2. Setting of model parameters

When ManEnt model is used to study the relationship between geographical distribution of vegetation and climatic factors, the applicability of the model should be evaluated first. The prediction ability, accuracy degree and error source of the model should be tested through the model evaluation process. The commonly used indicators for evaluating the fitness of a model include sensitivity, specificity, Kappa consistency test, and the Area Under ROC (AUC) (Li Fang, 1996; Silva et al., 2019; Sun Yu et al., 2014).

In order to study the influence of random training ratio of different current distribution points on the prediction accuracy of MaxEnt model, the geographic location data of alpine vegetation sample points in Hengduan Mountain area and the extracted environmental feature variable data were imported into MaxEnt3.4.1. Data of 65%, 70%, 75%, 80%, 85% and 90% of vegetation distribution sample points were randomly selected as training data sets to establish the model, and the remaining alpine vegetation sample points were used for model verification. Check Random Seed, other parameters remain unchanged, run 10 times in the pre-experiment, in order to make the AUC value obtained by simulation more stable (±0.001). AUC value is the area under the ROC curve provided by the model. Different values represent different levels of importance: 0.5-0.6, failing; 0.6-0.7, poor; 0.7-0.8, general; 0.8-0.9, good; 0.90-1.0, excellent (Phillips et al., 2006; Merow et al., 2014).

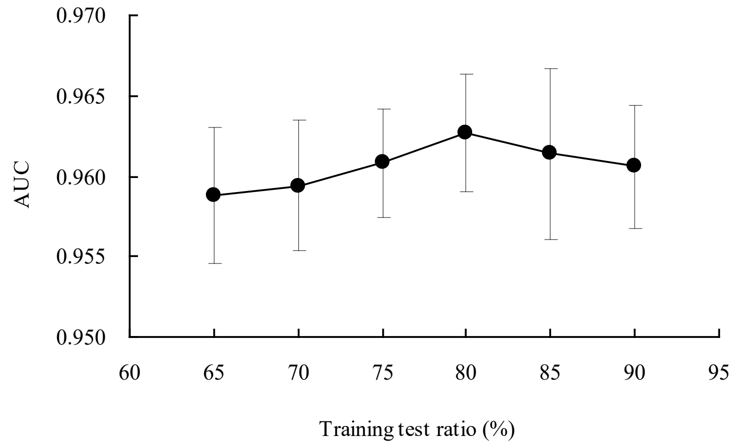
An evaluation criterion used in this paper is the standard deviation of AUC. In the process of data processing using the MaxEnt model, the operation and processing of each test percentage data are repeated for many times. The variance calculation formula of multiple AUC values calculated is as follows:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}\left(Xi - \overline{X}\right)^2}{n-1}} \tag{2}$$

In Formula (2), *SD* is the standard deviation of AUC, *n* is the number of repeats, *Xi* is the value of the ith AUC, and *X* is the average value of the *n* AUC. The higher the standard deviation of AUC is, the more discrete the obtained data is. In other words, the more unstable the simulated AUC value is, and the more inaccurate the model prediction data is. Conversely, the lower the standard deviation of AUC, the more stable the calculated AUC value and the more accurate the model prediction data obtained.

The 10 AUC values obtained from each training set were processed with mean square error by GraphPad Prism software, and the results in Figure 3 were obtained. It can be found that although the AUC value reaches the maximum value at 80% of the random training proportion, the variance value of AUC shows that the variance is the minimum at 75% of the random training proportion of the model, indicating that the AUC value is the most stable at this time and the data obtained by the model simulation is the most accurate. Therefore, the research results of this paper are based on the random selection of 75% training data proportion as the basis for the operation of the model.

Figure 3 AUC values and AUC variance of different training proportions.



3. Operation and validation of model

Unify boundary, coordinate system and raster size of all environment variables, convert raster into ASCII file that can be recognized by MaxEnt model, and save coordinates of alpine vegetation sample points in EXCEL software as.CSV format. The most critical step for MaxEnt model to study the relationship between vegetation geographical distribution and climate factors is to select the index of impact factors. The geographical distribution data of alpine vegetation in Hengduan Mountains and the extracted data of ecological climate variables were imported into MaxEnt3.4.1. 75% vegetation distribution sample points were randomly selected to establish the model, and the remaining vegetation distribution points were used for model verification. Check the random seeds, prediction chart and knife analysis, set the normalized multiplier, maximum iteration times, convergence threshold, maximum background number and other parameters, and run the MaxEnt model more than 10 times each time to get a more stable simulation result.

In this study, AUC value and Kappa value were selected as the criteria for judging the applicability and accuracy of MaxEnt model simulation. The evaluation criteria are shown in Table 1 (Swets et al., 1988).

Table 1 Relationship between AUC value and Kappa value and model accuracy

Accuracy	Excellent	Good	General	Poor	Extremely poor
AUC	0.9-1.0	0.8-0.9	0.7-0.8	0.6-0.7	0.5-0.6
Kappa	0.7-1.0	0.55-0.7	0.4-0.55	0.2-0.4	0.0-0.2

The evaluation results of the characteristic curve of the model simulating the geographical distribution of alpine vegetation in Hengduan Mountains were as follows: the average AUC was 0.934, reaching the excellent level. It shows that MaxEnt model can well simulate the potential suitable distribution area of alpine vegetation in Hengduan Mountains.

Second, Kappa was originally proposed by Cohen (1960) and has been widely used in the evaluation of classification accuracy (Tian Miao et al., 2012). Kappa consistency test was carried out between the potential distribution of alpine vegetation in the study area after the model simulation and the existing status point map, which could test the accuracy of the simulation on the whole. The obtained sample points were extracted from the numerical values of the simulation results from 1980 to 1999. If the sample points are in the suitable area, it means that the simulation results are consistent with the actual distribution. The Kappa coefficient is 0.627, which indicates that the model simulation accuracy reaches a good level.

The above two different test methods show that the MaxEnt model can well simulate the potential suitable distribution area of alpine vegetation in Hengduan Mountains.

4. Contribution rate of each climate variable

The geographical distribution of alpine vegetation is constrained by multiple environmental variables, each of which is irreplaceable, but some environmental variables are dominant in the geographical distribution of alpine vegetation. The percentage contribution rate of climatic factors to the geographical distribution of alpine vegetation given by MaxEnt model in the training process. In the process of model establishment and operation, cross validation is adopted to validate the model. First of all, the prediction effect of the potential distribution point of alpine vegetation simulated by the model and the importance of each environmental characteristic variable were tested by the AUC evaluation index in the model small tool method. This verification of AUC value was carried out in the form of sample points to verify the model. Scores of climatic factors in the small knife method to determine which climatic factors are indispensable to the simulation of climatic suitability geographical distribution of alpine vegetation (whether important or not). Secondly, ROC test is performed on environmental variables. Figure 17 shows the AUC value of climate factor in the small knife method to determine whether climate factor is essential (whether it is necessary).

5. Division of climate suitability

Different researchers have different classification methods for the results predicted by MaxEnt model, namely, the classification of suitability. Zhou et al. (2015) used the principle of different guarantee rate of climate resources to classify the suitability level. Tan et al. (2018) set the minimum suitability threshold as 0.10 and divided it between 0.10-0.30, 0.30-0.60 and 0.60-1.00. Wang et al. (2019) adopted the Manual classification method in the Reclass tool of ArcGIS software and divided it according to the suitability index, namely, 0.00-0.10 is the unsuitable area, 0.10-0.30 is the low suitable area, 0.30-0.50 is the marginal suitable area, 0.50-0.70 is the suitable area, and 0.70-1.00 is the most suitable area. Li et al. (2019) divided the suitability into three levels, in which the area with a potential distribution probability of 0.00-0.50 was the unsuitable area, the area with a potential distribution probability of 0.50-0.75 was the moderately suitable area, and the area with a potential distribution probability of 0.75-1.00 was the highly suitable area.

In this study, the above classification methods were tried respectively, and the results showed that according to the principle of climate guarantee rate, the proportion of the highly suitable area of alpine vegetation geographical distribution in the total area of the study area in the base period (1980-1999) was 10.04%. According to the classification method of Li et al. (2019), the area of moderate and highly suitable alpine vegetation in the geographical distribution accounted for 9.88% of the total area of the study area. According to the output results, ArcGIS10.2 software was used for the natural break point method (the data set was classified according to the discontinuous places in the data set), and the classification grades were 0.00-0.10, 0.10-0.28, 0.28-0.50, and 0.50-1.00, respectively. The height suitable area of alpine vegetation geographical distribution took up 8.94% of the total area of the study area. According to the vegetation type chart data (1:100,000) compiled by the famous vegetation ecologist Hou Xueyu in 2001, the alpine vegetation area in Hengduan Mountains is 25,652 km^2, about 8.2% of the total area of the study area, mainly distributed in Qinghai Province, Tibet Autonomous Region. Among them, the division method of natural break point is the closest to the actual geographical distribution range and area of alpine vegetation in the vegetation type map of the base period.

Then, the simulation results of habitat suitability were divided into four grades according to the natural break point method: 0-0.10 is unsuitable; 0.10-0.028 is low; 0.028-0.5 is moderate; and 0.50-1 is high. The climatic suitability grade distribution map of the geographical distribution of alpine vegetation in Hengduan Mountains obtain by ArcGIS.

The DEM layer was used to extract the altitude distribution of the alpine vegetation climate completely suitable area, and the change trend with altitude was calculated.

References

Li Wang-jun, FENG Tu, ZHOU Rui-wu, HE Bin, CUI Tao, DENG Fang-fang, PENG Mingchun. Maxent Model based on the potential Distribution prediction of natural Yellow Cedar forests in Guizhou Province [J]. Western Forestry Science, 2019, 48(3): 47-52.

Li Feng. Evaluation model of regional desertification degree [J]. Forestry Resources Management, 1996(3):62-66.

Merow C, Silander Jr J A. A comparison of Maxlike and Maxent for modelling species distributions[J]. Methods in Ecology and Evolution, 2014, 5(3): 215- 225.

Morales N S, Fernάndez I C, Baca-Gonzάlez V. MaxEnt’s parameter configuration and small samples: are we paying attention to recommendations?[J]. A systematic review. Peerj, 2017, 5(2): e3093.

Phillips S J, Anderson R P, Schapire R E. Maximum entropy modeling of species geographic distributions [J]. Ecological Modelling, 2006, 190(3/4): 231-259.

Renner I W, Warton D I. Equivalence of Maxent and poisson point process models for species distribution modeling in ecology[J]. Biometrics, 2013, 69(1): 274 – 281.

Silva C, Leiva F, Lastra J. Predicting the current and future suitable habitat distributions of the anchovy

(Engraulis ringens) using the Maxent model in the coastal areas off central-northern Chile[J]. Fisheries Oceanography. 2019,28(2): 171-182.

Sun Yu, Shi Ming-Chang, Peng Huan, Zhu Pei-Lin, Liu Si-Lin, Wu Shi-Lei, He Cheng, Chen Feng. Forest lighting fire forecasting for Daxing'anling Mountains based on MAXENT model[J]. The journal of applied ecology, 2014, 25(4):1100-1106.

Swets J A. Measuring the accuracy of diagnostic systems. Science, 1988, 240(4857) : 1285-1293.

Tian Miao, WANG Pengxin, YAN Tailai, LIU Chunhong. Modification of Kappa coefficient and its application in drought prediction accuracy and consistency evaluation [J]. Journal of Agricultural Engineering, 2012, 28(24): 1-7.

Zhang Weibo. Some Thoughts on the Current Situation of Social Network Communication based on entropy Theory [J]. News Knowledge, 2014(10): 21-23. (in Chinese with English abstract)