

Research Article

Potential Distribution Prediction of Medium Aroma Type of Flue-cured Tobacco Based on Ecological Niche Model

Shuangyang Xie, Naijia Guo and Hongbin Liu

College of Resources and Environment, Southwest University, Chongqing 400716, China

Abstract: In this study, six commonly used ecological niche models including genetic algorithm for rule-set production, support vector machine, maximum entropy and ecological niche factor analysis, DOMAIN and Bioclim were evaluated for predict potential suitable areas for flue-cured tobacco with medium aroma type. The ecological factors, such as topography and climate parameters, were collected from representative sites growing flue-cured tobacco with medium aroma type. Average receiver operating characteristic curve and area under the curve are calculated to assess the prediction accuracy of these models. The comparative results showed that maximum entropy model outperformed others. The potential distribution areas for producing flue-cured tobacco with medium aroma type are identified by the best performed model.

Keywords: Ecological niche model, habitat suitability, medium aroma type of flue-cured tobacco, model stability

INTRODUCTION

China is a big tobacco-producing and consuming country. The production of flue-cured tobacco in China accounted for 52% of global production (Liu, 2004). The tobacco-producing areas are distributed in 24 provinces include over 800 counties in China. The aroma types of flue-cured tobacco are classified into heavy, medium and light (Ding *et al.*, 1958). The total area producing flue-cured tobacco with medium aroma type is larger than that of heavy and light types across the country.

Ecological niche model is widely used in the risk analysis of harmful biological invasion (Wang, 2007; Sun and Liu, 2010), plant diseases and insect pests distribution prediction (Liu *et al.*, 2010), prediction and protect endangered species of wild fauna and flora distribution area (Wu and Lv, 2009), etc. Moreover, ecological niche model is also used to identify the suitable areas for growing particular crops or trees (Jones *et al.*, 2005).

Ecological conditions have significant effects on the growth and quality of flue-cured tobacco and hence the aroma types of tobacco (Yang *et al.*, 2014; Wu *et al.*, 2013a). The objective of the current work is to predict the potential suitable areas for producing flue-cured tobacco with medium aroma type based on ecological niche models. To achieve the aim, ecological factors including topography and climate were collected from the representative counties producing flue-cured tobacco with medium aroma type. Six ecological niche models are developed and compared using the collected

data and the best performed model is used to predict the potential areas which are suitable for planting medium aroma type of flue-cured tobacco. The results are expected to establish the ecological evaluation model or system of medium aroma type of flue-cured tobacco, plan the producing area and improve production technology.

MATERIALS AND METHODS

The tobacco data are obtained from the document of Tobacco company of China (2011, no.23), include 86 sites of light aroma type of flue-cured tobacco, 103 sites of medium aroma type of flue-cured tobacco and 63 sites of heavy aroma type of flue-cured tobacco (Fig. 1).

The digital elevation model and climate data are from <http://www.worldclim.org/>, the administrative map of China with the accuracy of 1:4000000 is from http://nfgis.nsd.gov.cn/nfgis/chinese/c_xz.htm and the solar radiation data is interpolated by Multiple Linear Regression (MLR) and Thin Plate Smoothing spline (TPS) (Wu *et al.*, 2013b). In order to compare the differences of the ecological environment of different types of tobacco, we select the average annual temperature, annual precipitation, annual total radiation of 30 years and elevation, slope of the known distribution area as the ecological environment factors (Table 1).

Six commonly used ecological niche models, namely, genetic algorithm for rule-set production

Corresponding Author: Hongbin Liu, College of Resources and Environment, Southwest University, Chongqing 400716, China

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

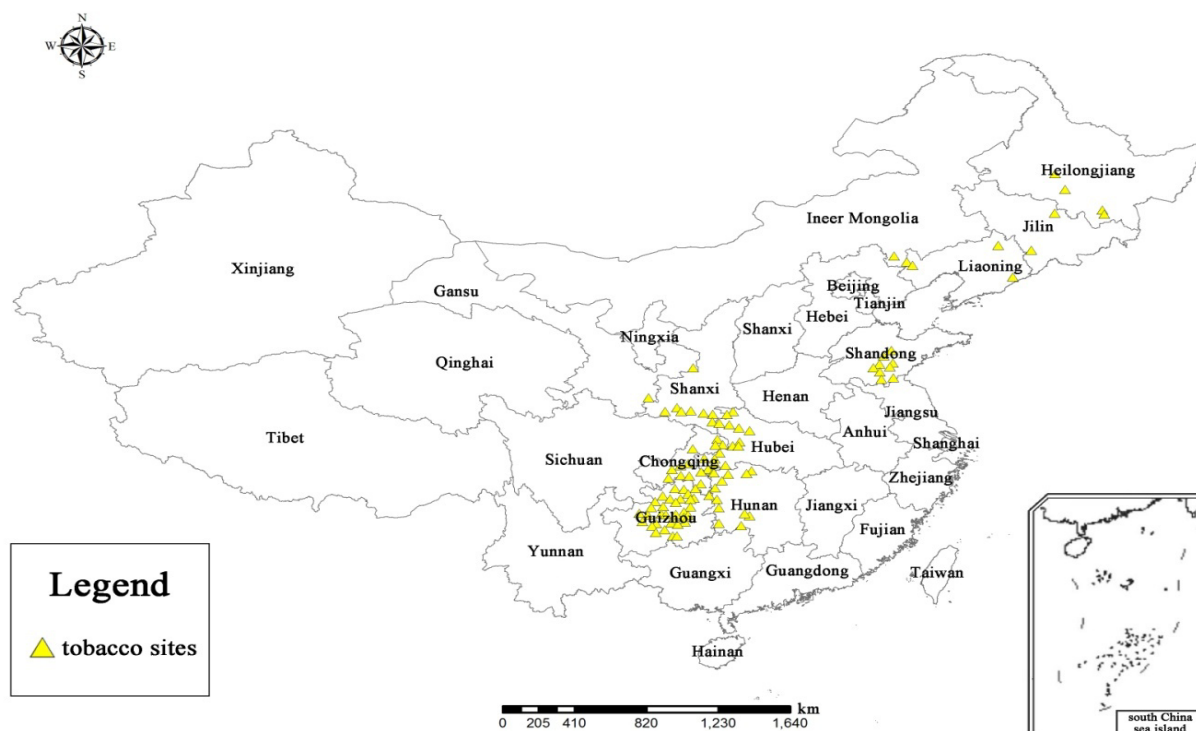


Fig. 1: Distribution of the three kinds of tobacco sites

Table 1: The description of source data

Data	Accuracy
Annual mean temperature	30"
Annual precipitation	30"
Annual solar radiation	30"
Elevation	30"
Slope	30"

(GARP) (Liu *et al.*, 2010), Support Vector Machine (SVM) (Drake *et al.*, 2006), maximum entropy (MaxEnt) (Phillips and Dudik, 2008) and Ecological Niche Factor Analysis (ENFA) (Mireia *et al.*, 2011), DOMAIN (Carpenter *et al.*, 1993) and Bioclim (Shao *et al.*, 2009) models are developed and compared in this study. In order to evaluate model performance, we make a random sampling on the distribution data of medium aroma type of flue-cured tobacco sites with different proportion (50, 65 and 80%), the data are then divided into training and validation sets. For each proportion, the process repeats 15 times. Therefore, 45 training and validation sets are created. Average Receiver Operating Characteristic (ROC) (Wang *et al.*, 2007) curve and Area Under the Curve (AUC) are calculated to test the prediction accuracy of these models. One-way analysis of variance with Least-Significant Difference (LSD) is applied to test the difference between the models' AUC at $p < 0.05$. Model stability and robustness is evaluated using different sample sizes (10, 20, 30, 40, 50, 60, 70, 80 and 90 samples). The optimal model is identified based on the stability and prediction accuracy. The optimal model will be used to identify the potential suitable areas for growing flue-cured tobacco with medium aroma type.

RESULTS AND DISCUSSION

Data overview: The statistics of the used ecological factors are shown in Table 2. Medium aroma type of tobacco are mainly distributed in the places where are warm, flat and with plentiful rainfall and sunshine. Temperature is mainly in the range of 13.96 to 15.40°C with a mean temperature of 14.68°C. Precipitation is mainly in the range of 1000 to 1104 mm with a mean precipitation of 1052 mm. Solar radiation is mainly in the range of 1454 to 1654 MJ/m² with a mean solar radiation of 1554 MJ/m². Elevation is mainly in the range of 447 to 621 m with a mean elevation is 549 m. Slope is mainly in the range of 2.18 to 3.55° with a mean value of 2.86°. It shows that medium aroma type of flue-cured tobacco has a relatively stable ecological niche.

Model performance: According to the proposed methods for evaluating the models performance, three average values of AUC for each model under different sampling plans are obtained and shown in Fig. 2 to 4. Under the 50% proportion, the average values of AUC of the six models are in order of Maxent>Domain>GARP>SVM>ENFA>Bioclim. Significant differences in AUC are found between the first four (Maxent, Domain, GARP and SVM) and the least two models (ENFA and Bioclim). Under the 65% proportion, the average values of AUC of the six

models are in order of Maxent>Domain>SVM > GARP >ENFA>Bioclim. Significant differences in AUC are found between the first four (Maxent, Domain, SVM and GARP) and the least two models (ENFA and Bioclim). Under the 80% proportion, the results of

AUC and ANOVA are similar to that of 65% proportion. This indicates that Domain, Maxent, SVM and GARP perform better than ENFA and Bioclim models. These four models can be used as the alternative models of the optimal model.

Table 2: The statistic description of the ecological factors

Ecological factor	Mean	S.D.	95% confidence interval for mean		Min.	Max.
			Lower bound	Upper bound		
Temperature (°C)	14.68	3.66	13.96	15.40	2.5	18.7
Precipitation(mm)	1052	267.54	1000	1104	447	1473
Solar radiation (MJ/m ²)	1554	510.46	1454	1654	1053	2829
Elevation (m)	549	367.98	447	621	46	1738
Slope (°)	2.86	3.49	2.18	3.55	0	19

S.D.: Standard deviation; Min.: Minimum; Max.: Maximum

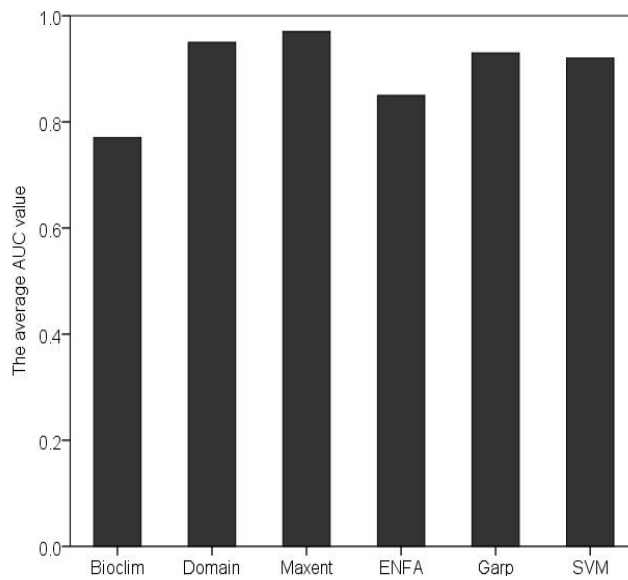


Fig. 2: The average AUC values of different models for proportion of 50%

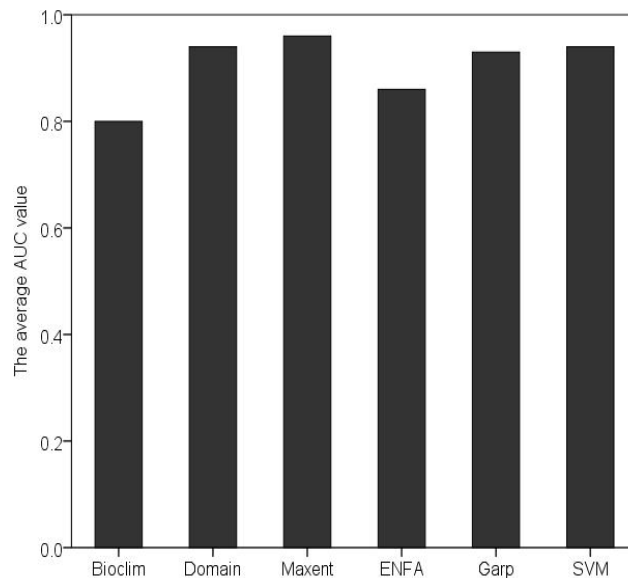


Fig. 3: The average AUC values of different models for proportion of 65%

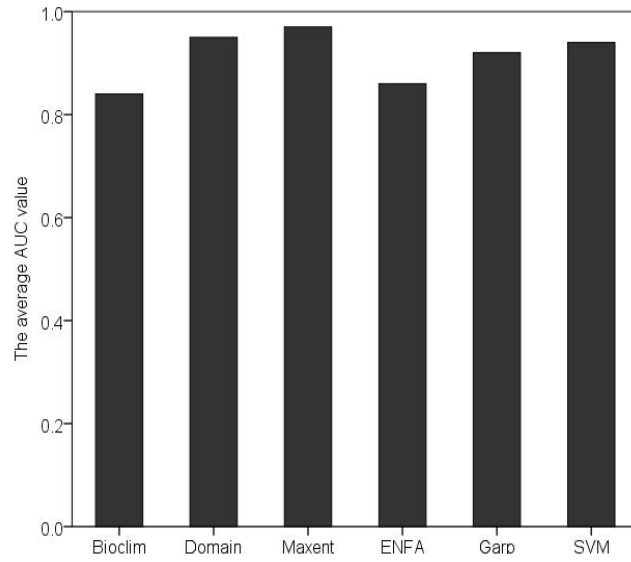


Fig. 4: The average AUC values of different models for proportion of 80%

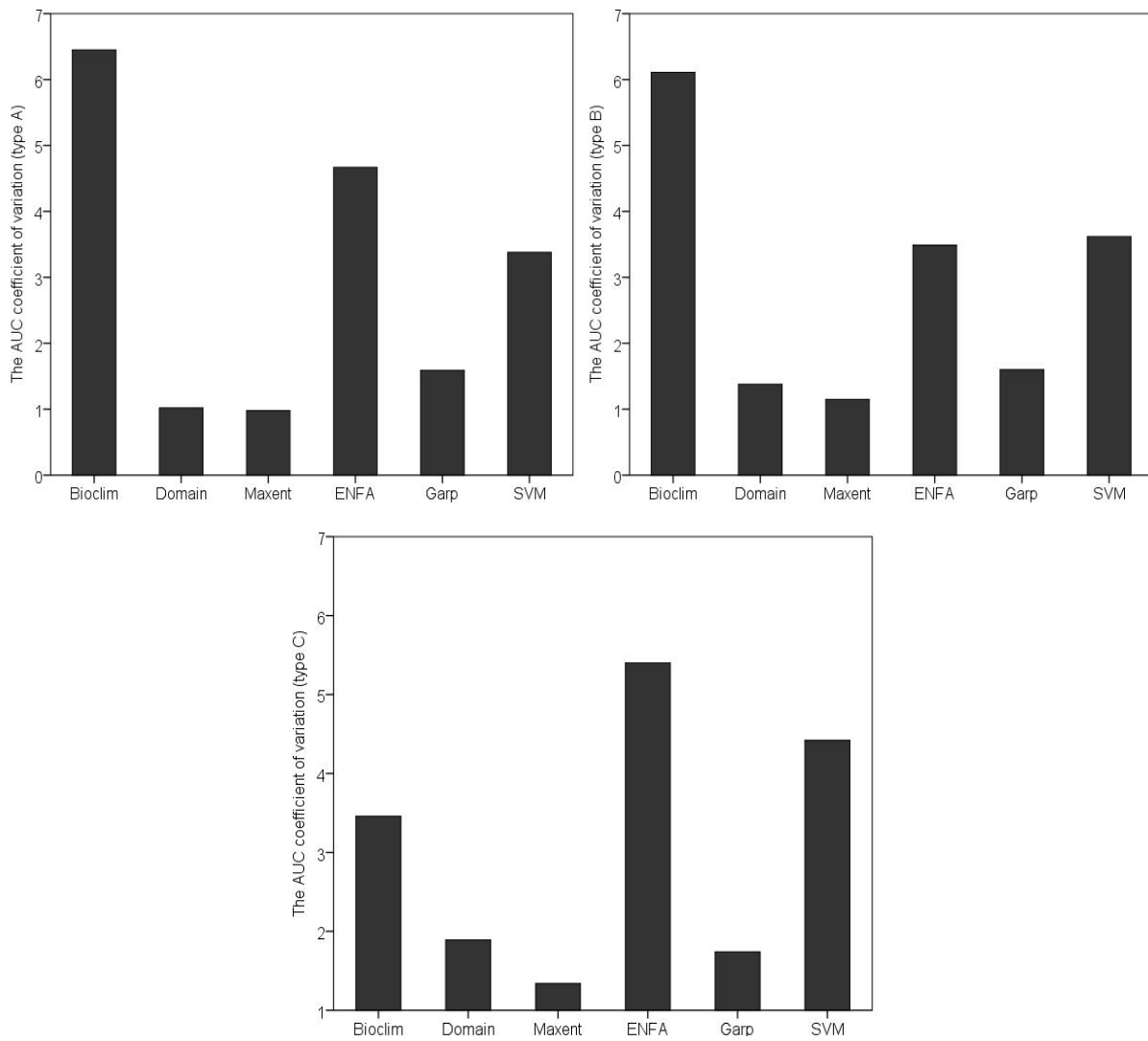


Fig. 5: Coefficients of variation of AUC for the six models under different proportions

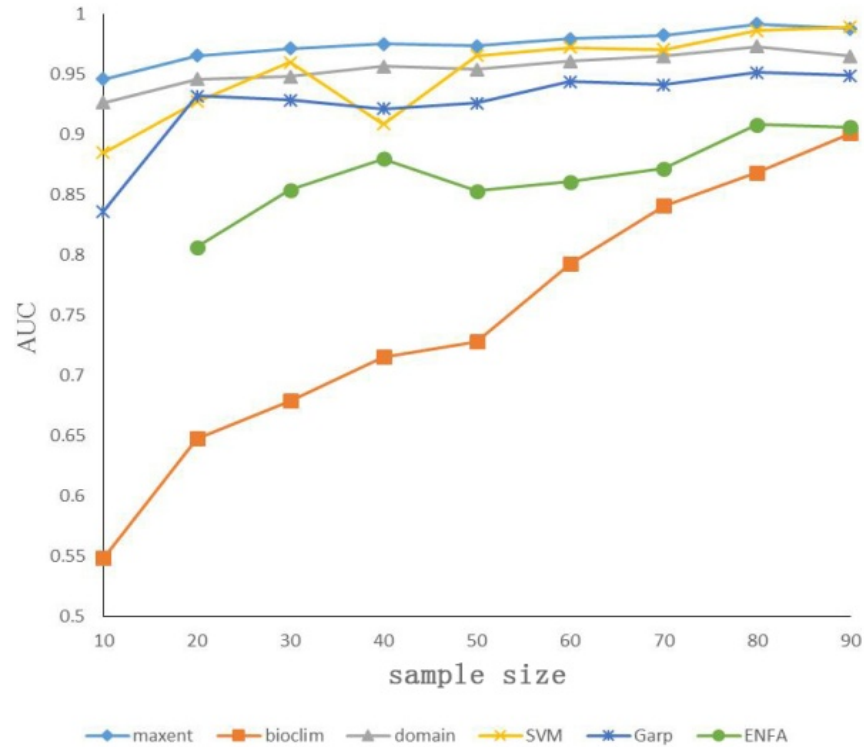


Fig. 6: The change of AUC with sample numbers for different models

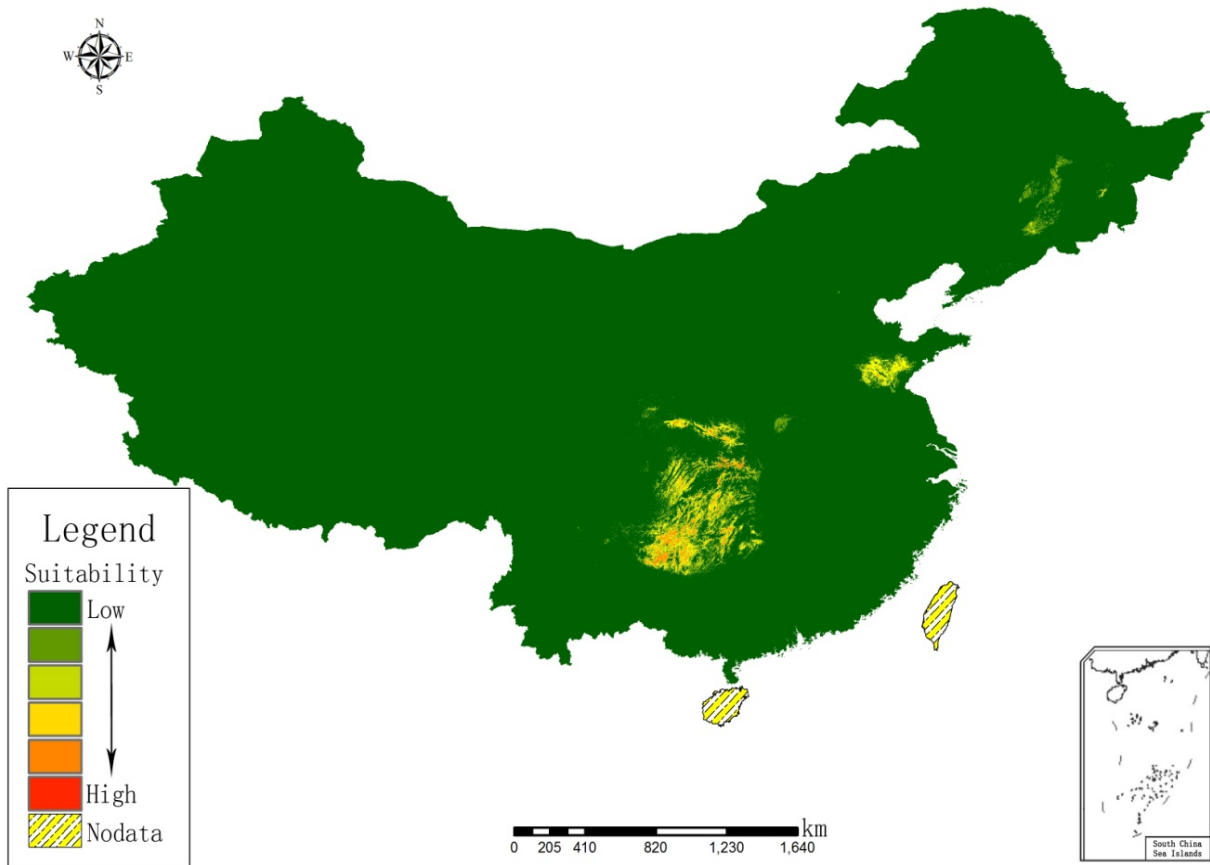


Fig. 7: The potential distribution map of medium aroma type of flue-cured tobacco

Coefficients of variation of AUC for the six models under the three proportions are also calculated and shown in Fig. 5. It is assumed that models with lower coefficient of variation of AUC are more stable. Results show that Domain, GARP and MaxEnt models have lower coefficient of variation under the three proportions.

A good model should have good stability and strong robustness, which could be able to provide accurate forecast under different sample sizes. Here, the relationships between AUC and sample sizes (10, 20, 30, 40, 50, 60, 70, 80 and 90) for the six models are also compared and given in Fig. 6. The values of AUC of the six models increase with the increasing in sample sizes. However, the amplitudes of the changes differ with models. For example, the largest change exists in Bioclim model, which increases from 0.55 to 0.88. For GARP model, the value of AUC varies from 0.84 to 0.93 when the number of samples changes from 10 to 20. But the change becomes small while the sample sizes are greater than 20. The results show that Maxent and Domain models are more stable and robustness than others.

The optimal model and prediction map: Maxent which has the largest value of AUC and the stable relationship between AUC and sample sizes is the optimal model and could be applied to predict the suitable areas for planting flue-cured tobacco with medium aroma type. The spatial distribution of the suitable areas is illustrated in Fig. 7.

The potential suitable areas predicted by the best model are mainly distributed in Guizhou, Chongqing, south Gansu, south Shanxi, west Hubei, west Hunan, south Shandong, central Jilin and minority in south Heilongjiang and Henan. The areas with higher suitability are located in southwestern China, especially Guizhou province which is a well-known flue-cured tobacco with medium aroma type production region. Some areas with lower suitability located in Henan province are mixed with their well-known aroma types of flue-cured tobacco. This indicates that more factors should be considered to improve the prediction accuracy.

CONCLUSION

This study evaluates six ecological niche models for predicting potential suitable areas for planting flue-cured tobacco with medium aroma type. Based on the proposed sampling methods, maximum entropy model outperformed rule-set production, support vector machine, ecological niche factor analysis, OMAIN and Bioclim models. The predicted potential suitable areas by the best model largely agree with their well-known aroma type. Future work will be needed to consider more ecological factors such as soil types and soil nutrients. And identify the effect of each factor will be the research emphasis in the future.

ACKNOWLEDGMENT

This research was supported by the science and technology project of the China Tobacco Corporation of Chongqing (NY20130301070006).

REFERENCES

- Carpenter, G., A.N. Gillison and J. Winter, 1993. DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals. *Biodivers. Conserv.*, 2(6): 667-680.
- Ding, R.K., C.H. Wang and Z.Q. Zhu, 1958. *Cigarette Technology*. Food Industry Press, Beijing.
- Drake, J.M., C. Randin and A. Guisan, 2006. Modelling ecological niches with support vector machines. *J. Appl. Ecol.*, 43: 424-432.
- Jones, P.G., W. Diaz and J.H. Cock, 2005. Homologue: A Computer System for Identifying Similar Environments throughout the Tropical World. Version Beta a.0. Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.
- Liu, J.H., X.Z. Xiong, L.Y. Yang, Z.P. Xiong and Y.Z. Pan, 2010. Analysis on suitable distribution area of oriental fruit fly (*Bactrocera dorsalis*) based on GARP ecological niche model. *Plant Dis. Pests*, 1(6): 34-38.
- Liu, T.N., 2004. *Tobacco Economy and Tobacco Control*. Economic Science Press, Beijing.
- Mireia, V., B. Angel, C. Guillem, G. Ibon and M.G. Joxe, 2011. Modelling suitable estuarine habitats for *Zostera noltii*, using ecological niche factor analysis and bathymetric lidar. *Estuar. Coast. Shelf S.*, 94: 144-154.
- Phillips, S.J. and M. Dudik, 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2): 161-175.
- Shao, H., J.Q. Tian, K. Guo and J.X. Sun, 2009. Effects of sample size and species traits on performance of BIOCLIM in predicting geographical distribution of tree species—a case study with 12 deciduous Quercus species indigenous to China. *Chinese J. Plant Ecol.*, 33(5): 870-877.
- Sun, W.T. and Y.T. Liu, 2010. Research progress of risk analysis of biological invasion. *Chinese Agric. Sci. Bull.*, 26(7): 233-236.
- Wang, Y.S., 2007. Application of niche models in the risk assessment of invasive alien species. Ph.D. Thesis, Hunan Agricultural University, Changsha.
- Wang, Y.S., B.Y. Xie, F.H. Wan *et al.*, 2007. Application of ROC curve analysis in evaluating the performance of alien species' potential distribution models. *Biodivers. Sci.*, 15(4): 365-372.
- Wu, J.G. and J.J. Lv, 2009. Potential effects of climate change on the distributions of Yunnan snub-nosed monkey (*Pygathrix Bieti*) in China. *J. Meteorol. Environ.*, 25(6): 1-10.

- Wu, W., X.P. Tang, C. Yang, H.B. Liu and N.J. Guo, 2013a. Investigation of ecological factors controlling quality of flue-cured tobacco (*Nicotiana tabacum* L.) using classification methods. *Ecol. Inform.*, 16: 53-61.
- Wu, W., X.P. Tang, C. Yang, H.B. Liu and N.J. Guo, 2013b. Spatial estimation of monthly mean daily sunshine hours and solar radiation across mainland China. *Renew. Energ.*, 57: 546-553.
- Yang, C., W. Wu, S.C. Wu, H.B. Liu and Q. Peng, 2014. Aroma types of flue-cured tobacco in China: Spatial distribution and association with climatic factors. *Theor. Appl. Climatol.*, 115(3): 541-549.