UDC 614.84:253(597)

APPLYING MULTIVARIATE STATISTICAL METHODS FOR PREDICTING PINUS FOREST FIRE DANGER AT BIDOUP-NUI BA NATIONAL PARK

Le Van Huong[1], Nguyen Ngoc Kieng[2], Nguyen Dang Hoi[3], Dang Hung Cuong[3,4]

[1]*Bidoup-Nui Ba National Park, Danhim Village, Lacduongdictrist, Lamdong Province, Vietnam*
[2]*Hochiminh University of agriculture and forestry, Ho Chi Minh, Vietnam*
[3]*Vietnam – Russia Tropical Center, Ha Noi, Vietnam*
[4]*Southern Federal University, Russia, Rostov-on-Don, Russia*
e-mail: lvhuongdl@gmail.com

The paper presents results of applying multivariate statistical methods (***CCA***: canonical correlation analysis and ***DFA***: discriminant function analysis) for determining canonical correlation between a set of variables {**T, H, m_1, K**} and a set of variables {**Pc, Tc**} (***T***: temperature, ***H***: relative humidity, ***m_1***: mass of dry fuels, ***K***: burning coefficient, ***K = m_1/M***, with ***M***: total mass of fire fuels, ***Pc***: % burned fuels and ***Tc***: burningtime) as well as through results of discriminant function analysis DFA to set up models of predicting forest fire danger at Bidoup - Nui Ba National Park. From research data in November, December, January, February and March in the period of 2015-2017 from 340 sampling plots (each 2mx2m), at Bidoup - Nui Ba National Park, we carry on data processing on Excel (calculating) and Statgraphics (multivariate statistical methods: ***CCA&DFA***). Three results were revealed from our analysis: ***(i)*** Canonical correlation between a set of variables {***T, H, m_1, K***} and a set of variables {***Pc, Tc***} is highly significant (R = 0.675581 & P = $3.17*10^{-58} << 0.05$); therefore, we can use a set of variables {***T, H, m_1, K***} in models of predicting forest fire danger, ***(ii)*** Coefficients of standardized & unstandardized canonical discriminant functions (***SCDF &UCDF***) and Fisher classification function (***FCF***) are determined, ***(iii)*** Setting up two models of predicting forest fire danger (Mahalanobis distance model & Fisher classification function model).

Key words: Bidoup-Nui Ba, canonical correlation analysis, discriminant function analysis, Fisher classification function, models of predicting forest fire danger.

Introduction

Forest fire is common in both tropical and temperate regions, especially during drought period. Forest fire depends on many different factors including temperature, humidity, the status of vegetation, the layer of fallen leaves. Of these factors, temperature (***T***) and relative humidity (***H***) were often used to establish formulas to access the risk of forest fire (Angstrom, 1942; Chandler, 1983; Viney, 1991; Cheney & Sullivan, 1997; Sharples, 2009).

However, the practice shows that the characteristics of fire materials in the forest also play an important role. Our researches in Bidoup – Nui Ba National Park on the relationship between temperature (T) and relative humidity (H) to forest fire risk also showed a negative correlation [$H = (14.2545 - 0.261584 * T)^2$, with a nonlinear correlation coefficient R = 0.784535 and probability level significance $P = 3.89* 10^{-72} << 0.05$]. It means that, when the temperature rises and the humidity of the forest environment is low, it will increase the evaporation of the fire material, making the material dry faster, leading to a higher risk of forest fires. Regarding the status of vegetation and the risk of forest fires, Le Van Huong (2012) studied and proposed the burning coefficient ***K*** and the mass of dry fuels m_1 playing an important role in the forecasting model of forest fire risk. The coefficient ***K*** (with $K = m_1$ / total mass of fire fuels ***M***; $M = m_1 + m_2$, where m_2 is the mass of fresh materials) and the mass of dry fuels m_1 shows a high correlation with the burned fuels ***Pc*** [$Pc = \exp[b_0 + b_1*K + b_2*m_1 + b_3*\ln K + b_4*\ln(m_1) + b_5*\sqrt{K} + b_6*\sqrt{m_1} + b_7/K + b_8/m_1 + b_9*\sqrt{K^3} + b_{10}*m_1^2 + b_{11}/K^2 + b_{12}/m_1^2]$ with multiple correlation coefficient R = 0.98 and probability level significance $P_{model} = 1.39 * 10^{-9} << 0.05$ and $P_{bi} < 0.05$; i = 1, 2, ..., 12. This is understandable, a large coefficient K is only a necessary condition but not enough for a fire to develop, and a sufficient condition for a large fire to grow, the mass of dry fuelsm_1 must be large enough. To clarify the dependence of forest fire on factors: temperature, humidity and state of vegetation, the article applies multivariate statistical methods to analyze data with a combination of all four factors for the forecasting model of pine forest fire risk in Bidoup - Nui Ba National Park.

Data Bases and Research Methods

Research data was collected in November, December, January, February and March in the period of 2015-2017 from 340 sample plots with size of 2m x 2m (Table 1). At each plot measured temperature T (°C) and humidity H (%). On the surface of each plot, classify the dry and flammable fuels (m_1, kg) or fresh and inflammable materials (m_2, kg). However, the mass of fresh materials can turn dry fuels as the fire develops. Therefore, the total mass $M = m_1 + m_2$ contains the potential to grow into a great fire later, while m_1 and $K = m_1/M$ are the factors for the fire to form and develop as mentioned above.However, at the time of prediction before the fire has formed, M can be calculated from $m1$ and K[$M = m_1/K$, and $m_2 = M- m_1 = m_1/ (K - m_1)$], so it is not necessary set M and m_2 in the forecasting model. After determining the masses of m_1 and m_2, mix the amount of dry fuels and fresh materials together and burn to collect the data Pc (% % burned fuels) and Tc (burning time, minutes). Burning time Tc denotes the rate of burning, meaning the shorter the burning time, the more intense the fire rate and vice versa. Thus, the variables Pc and Tc that describe the results of the fire should be called as a set of dependent variables. The variables T, H, m_1 and K which are the causes of the fire should be called as a set of independent variables.

Table 1.

Time, location, forest type of 340 sample plots conducting data collection in Bidoup - Nui Ba National Park

No.	Time	Location Sub-area	Forest type	Quantum of sample plots
1	20/11/2015	145A	Natural forest	3
2	29/12/2015	27	Natural forest	10
3	30, 31/12/2015	26	Plantation forest in 2002	40
4	16, 17/01/2016	103	Plantation forest in 2002	30
5	24, 25/01/2016	76	Plantation forest in 1996	30
6	30/01/2016	145A	Natural forest	27
7	01/02/2016	59	Plantation forest in 1998	30
8	28/02/2016	80	Plantation forest in 1996	15
9	5,6/03/2016	96C	Plantation forest in 1997	35
10	22, 23/03/2016	124	Plantation forest in 1991	30
11	29, 30/11/2016	100	Plantation forest in 2011	45
12	18/02/2017	102A	Plantation forest in 1999	15
13	19/02/2017	75B	Plantation forest in 1999	15
14	19/02/2017	93	Plantation forest in 1998	15
Total				**340**

Note: Collect data to experimentally verify the fire risk forecast models in Bidoup - Nui Ba National Park

In order to have basis data for calculation in the forest fire forecast at a certain time in Bidoup – Nui Ba National Park, sample plots (size 2m x 2m) were established. The T,H, m_1 and M data, with K calculated by the formula $K = m_1/M$ were collected in the sample plots. Then we took the average values of the T, H, m_1 and K variables from those sample plots and calculated Mahalanobis multivariate distances based on the correlation discriminant functions CDF and the values of the Fisher's classification functions (FCF). From these calculation results, a fire risk forecast for Bidoup – Nui Ba National Park will be determined. For testing, we collected a database of variables included in the calculations in the models at different times during the dry season in Bidoup – Nui Ba National Park, the results are summarized in the Table 2.

Table 2.

The database (T, H, m_1, K) collected at different times in the dry season in Bidoup – Nui Ba National Park

Time	T	H	m_1	K
November	27.00	51.40	0.85	0.141
December	29.26	55.14	1.58	0.652
January	25.04	73.66	2.40	0.542
February	25.00	53.40	5.93	0.85
March	30.24	28.12	1.54	0.522

*Note: **T** and **H** are the measurements of temperature (^oC) and humidity (%) on the surface of forest land at the time of the survey; **K** = m_1/M, where m_1 is the weight of dry fuels per sample plot and **M** is the total weight of both dry fuels and fresh materials per sample plot.*

This input database (including variables T, H, m1 and K, at different times during the dry season from November to March) is put into actual testing of models (model Mahalanobis and Fisher's classification function model). Results will be presented in the following sections.

Data were processed using Excel (calculation) and Statgraphics software (multivariate statistical methods such as CCA: canonical correlation analysis, and DFA: discriminant function analysis). The basic contents of the canonical correlation analysis (CCA) method and discriminant function analysis (DFA) method are as follows:

(i) Unlike the problem of data analysis on correlation and regression, using univariate statistical methods is data collected consisting of one or more a set of independent variables x_i but only a single dependent variable ***y***. In the field of using multivariate statistical methods to perform canonical correlation analysis (CCA), data collection consists of a set of independent variables x_i and a set of dependent variables y_i. The object of canonical correlation analysis is to convert the original X and Y variable sets (in this article ***X = {T, H, m1, K}*** and ***Y*** = {***Pc, Tc***}) into pairs canonical variable (***Ui, Vi***) as its representative. Each pair of canonical variables(***U1, V1***)**, ...,** (***Ur, Vr***) are completely independent of each other, with pairs of canonical variables (U1, V1) being the best, most important and reliable representatives, followed by (***U2, V2***), ..., (***Ur, Vr***) (Wilks, 2011). As a result of the canonical correlation analysis, we can evaluate the correlation of the X and Y variables sets by the correlation coefficients ***R, R***$=\sqrt{\lambda}$ with λ is eigenvalue, and the probability of significance ***P***, show visualization on the ordination map (Wilks, 2011) of research objects (in this article, the results are derived from the data of 340 sample plots).

(ii) The purpose of the discriminant function analysis (DFA) method (*Poulsen, French, 2004*) is based on the data of x_i variables to separate the research objects into different groups. For example, in this article, there are 4 variables x_i (T, H, m1, K) that separate the study objects (340 sample plots) into 5 levels of forest fire risk.And if there are groups in practice, we need to consider whether these groups are different from a statistical stand point.The results of the discriminant function analysis give us the following major products:

(a) Canonical discriminant functions (CDF), together with standard correlation coefficients ***R***, ***R*** $=\sqrt{\frac{\lambda}{1+\lambda}}$ with λ being eigenvalueand the probability probabilities ***P*** corresponding to the functionsare used in many fields (Dillon, 1984; Cruz-Castillo, 1994; Edzard, 2002; Matthew, 2010).From thesecanonical discriminant functions,we can establish a predictive model for forest fire risk based on the multivariate distance Mahalanobis (Mahalanobis, 1930; Rao, 1973a, 1973b; Gupta, 1998; Ghosh, 1998; McLachlan, 1999).

(b) Fisher's classification functions were first introduced by Fisher (Srivastava, 2002; Anja Hashagen, 2009) or Fisher’s linear discriminant functions. We can establish a model for forecasting the risk of forest fire based on this set of Fisher’s classification functions.

Results and Discussion

The results of the standard correlation analysis between the independent and dependent variable sets

In order to prove that the independent variable {***T, H, m_1, K***} and the dependent variable {***Pc, Tc***} have statistically reliable correlations, the canonical correlation analysis (CCA) method has been appliedfor processing research data.

The results of canonical correlation analysis between independent variables {***T, H, m_1, K***} and dependent variables {***Pc, Tc***} on Statgraphics are as follows:

X = - 0.102 * ***T'***+ 0.501 * ***H'***+ 0.417 * ***m_1'***- 0.988 * ***K'*** (for ***T'***, ***H'***, ***m_1'***and ***K'*** are normalized variables from the variables ***T, H, m_1*** and ***K***) is the positioning function on the horizontal axis of the independent variable set {***T, H, m_1, K***} (set 1).

Y = 0.527 * ***Tc***' - 0.839 * ***Pc'*** (with ***Tc'***and ***Pc'*** are the standardized variables from ***Tc*** and ***Pc*** variables) is the positioning function on the vertical axis of the dependent variable set {***Pc, Tc***} (set 2).

Convert ***T, H, m_1*** and ***K*** variables into normalized variables ***T'***, ***H'***, ***m_1'***and ***K'*** as follows: ***$T'=(T-m_T)/S_T$***, where ***m_T*** is the average of ***T***, and ***S_T*** is the standard deviation for ***T***; similar to the standardized variables ***H'***, ***m_1'*** and ***K'***.

The standard correlation coefficient between the independent variable {***T, H, m_1, K***} and the dependent variable set {***Pc, Tc***} is ***R*** = 0.675, with probability significance ***P*** = $3.17*10^{-58} << 0.05$ (fig 1).

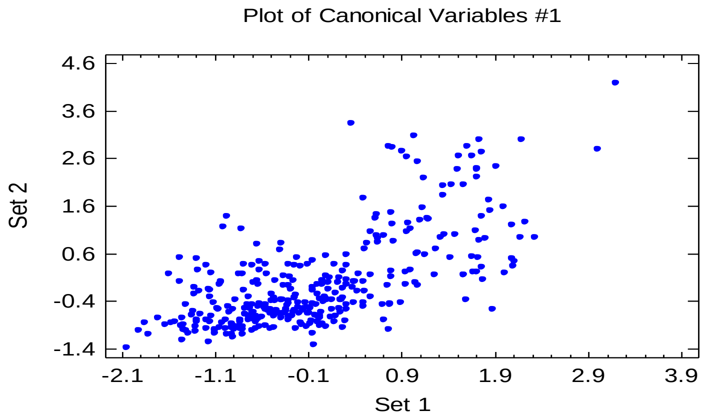

Figure 1. The results show the standard correlation between the independent variable set X = {T, H, m_1, K} and the dependent variable Y = {Tc, Pc}

From the results of the canonical correlation analysis between the independent variable {T, H, m1, K} and the dependent variable {Pc, Tc}, prove that the correlation between them is high and significant in terms of statistical. Therefore, it is very appropriate and objective to include variables ***T***, ***H***, ***m_1*** and ***K*** into the forest fire prediction model.

The result sets up canonical discriminant function CDF and the Fisher's classification function FCF

The result sets thecanonical discriminant function CDF

CDF canonical discriminant function, also known as DF discriminant function, is the most established functions with unstandardized variables, deduced from the discriminant function with the standardized variables. Applying DFAdiscriminant function analysismethod on Statgracphics to process research data of 340 sample plots in Bidoup - Nui Ba National Park, the results of establishing the standard calibration functions are as follows (CDF1, CDF2, CDF3 & figure 2):

CDF1 = -17.395+0.164 * ***T***+ 0.197****H***+0.129****m_1***+4.083****K***

With the canonical correlation coefficient ***R*** = 0.926 and the probability significant ***P*** = $2.09 * 10^{-259} << 0.05$

CDF2 = $-3.269 + 0.0414*\boldsymbol{T} - 0.0335*\boldsymbol{H} + 0.777*\boldsymbol{m_1} + 3.795*\boldsymbol{K}$

With the canonical correlation coefficient $\boldsymbol{R} = 0.8367$ and the probability significant $\boldsymbol{P} = 7{,}82*10^{-126} \ll 0.05$

CDF3 = $-13.162 + 0.3239*\boldsymbol{T} + 0.049*\boldsymbol{H} - 1.033*\boldsymbol{m_1} + 6.721*\boldsymbol{K}$

With the canonical correlation coefficient $\boldsymbol{R} = 0.681$and the probability significant $\boldsymbol{P} = 5{,}55*10^{-44} \ll 0.05$

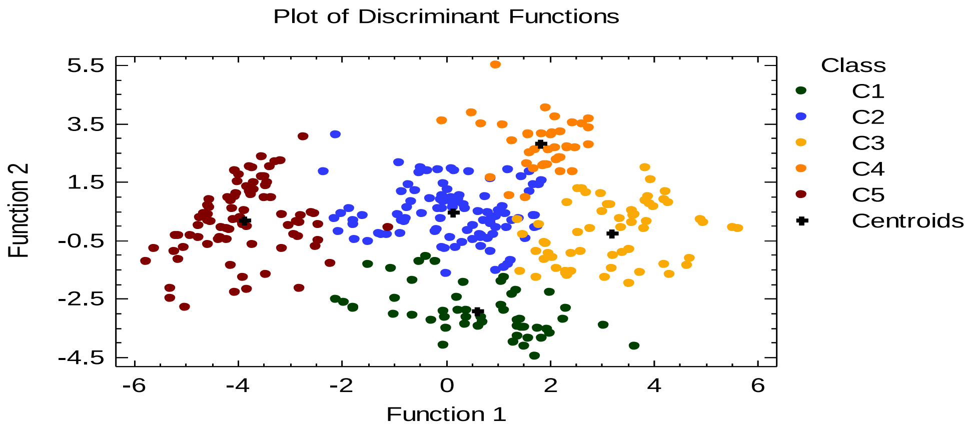

Figure 2. Ordination map and classification of forest fire risk class in Bidoup-Nui Ba National Park

The analysis results show that all three functions CDF1, CDF2 & CDF3 are very statistically meaningful ($P \ll 0.05$) accounting for a very high proportion (99.99%). And the CDF4 function is not statistically significant because the standard correlation coefficient $R = 0.0256$ is very low, with a probability that is not statistically significant $P = 0.638 \gg 0.05$, accounting for only 0.01%, should be excluded from the calculation. Therefore, we only use 3 functions of CDF1, CDF2 & CDF3 in calculating the risk of forest fire in Bidoup-Nui Ba National Park.

Result of sets the Fisher's classification functions

The Fisher's classification functions have also been established based on the results of the discriminant function analysisas follows:

FCF1 = $-389.679 + 15.384*\boldsymbol{T} + 5.5073*\boldsymbol{H} + 0.147*\boldsymbol{m_1} + 124.026*\boldsymbol{K}$ (forest fire classification function class 1, C1).

FCF2 = $-414.731 + 16.096*\boldsymbol{T} + 5.399*\boldsymbol{H} + 0.626*\boldsymbol{m_1} + 148.486*\boldsymbol{K}$ (forest fire classification function class 2, C2).

FCF3 = $-461.368 + 16.361*\boldsymbol{T} + 5.997*\boldsymbol{H} + 1.251*\boldsymbol{m_1} + 153.093*\boldsymbol{K}$ (forest fire classification function class 3, C3).

FCF4 = $-422.074 + 15.583*\boldsymbol{T} + 5.521*\boldsymbol{H} + 5.542*\boldsymbol{m_1} + 145.743*\boldsymbol{K}$ (forest fire classification function class 4, C4).

FCF5 = $-334.679 + 15.029*\boldsymbol{T} + 4.553*\boldsymbol{H} + 1.255*\boldsymbol{m_1} + 122.319*\boldsymbol{K}$ (forest fire classification function class 5, C5).

These Fisher classification functions are used to determine the fire risk class when we provide the input data (T, H, m1 and K).The forest fire risk class is forecasted with the calculated value of the Fisher $\boldsymbol{FCF_i}$ classification function ($\boldsymbol{i}$ = 1, 2, 3, 4, 5) corresponding to the forest fire risk class$\boldsymbol{C_i}$ (i = 1, 2, 3, 4, 5) is the largest.

Forecast of forest fire risk is based on the results of the discriminant function analysis

Forecast of forest fire risk is based on the Fisher classification function

The forest fire risk class is forecasted with the calculated value of the Fisher's classification function $\boldsymbol{FCF_i}$ ($\boldsymbol{i}$ = 1, 2, 3, 4, 5) corresponding to the forest fire risk class$\boldsymbol{C_i}$ ($\boldsymbol{i}$ = 1, 2, 3, 4, 5) is the largest. Used the database of variables included in Table 1 and based on 5

Fisher's classification functions in the prediction model to calculate classification scores, the results are in Table 3.

Table 3.
Result of forecasting the risk of forest fire at different times in the dry season in Bidoup-Nui Ba National Park based on the Fisher classification function

Time	Variables included				Maximum value of the Fisher's function	Forecast class
	T (°C)	***H*** (%)	m_1 (kg/4m^2)	***K***		
November	27	51.4	0.85	0.141	326.412	C1
December	29.26	55.14	1.58	0.652	451.767	C2
January	25.04	73.66	2.4	0.542	476.077	C3
February	25	53.4	5.93	0.850	419.042	C4
March	30.24	28.12	1.54	0.522	313.6420	C5

From the results of calculating the classification scores of the Fisher's classification functions shown in Table 3, we can give the following conclusions:

- The calculation results from the forest fire risk prediction model based on the Fisher's classification function also give the same results from the model based on the Mahalanobis distance.
- The calculation process in the forest fire risk prediction model based on the Fisher's classification functions is much simpler than the calculation based on Mahalanobis distance.

Forecast of forest fire risk based on Mahalanobis distance

As mentioned above, based on the established correlation discriminant functions CDF we can calculate Mahalanobis multivariate distances.The Mahalanobis multivariate distance D^2in discriminant functionsanalysis is the square Euclidean multivariate distance from the positioning coordinates of the input variable {T, H, m1, K} to the positioning coordinates of the center of forest fire class (C1, C2, C3, C4, C5).The risk of forest fire is forecasted with the shortest Mahalanobis distance.Using the database of variables in Table 1, based on correlation discriminant functionsCDF to calculate the multivariate distance Mahalanobis will result in Table 4.

Table 4.
Result of forecasting the risk of forest fire at different times in the dry season in Bidoup-Nui Ba National Park based on the Mahalanobis multivariate range

Time	Variables included				The shortest distance Mahalanobis	Class
	T(°C)	***H*** (%)	m_1 (kg/4m^2)	***K***		
November	27	51.4	0.85	0.141346	8.138	C1
December	29.26	55.14	1.58	0.652034	2.042	C2
January	25.04	73.66	2.4	0.542221	1.023	C3
February	25	53.4	5.93	0.85001	2.4656	C4
March	30.24	28.12	1.54	0.52215	0.433	C5

From the results of the Mahalanobis multivariate distance calculation shown in Table 4, we can draw the following conclusions:

- With data collected in November, the model predicted and calculated as the result of the forest fire risk class C1.At class C1, there is very little chance of forest fire.This result is also consistent with the fact that at the beginning of the dry season, because the amount of water accumulated in the dry material is still high, the amount of dry material is low, so the forest is very difficult to burn. At this time, if proactive burning is conducted to prevent forest fires, it will not be much effective, because the dry material layer of the forest will be very difficult to catch fire or burn very little, insignificantly.
- With the data collected in December, the model predicted and calculated as a result of

C2 forest fire riskclass. There is little chance of forest fires. This is the time to start using proactive combustion solutions for forest fire prevention.

- With data collected in January, the model predicted and calculated as the result of C3 forest fire risk level. With thisclass, there is a possibility of forest fire. This may be the best time for the proactive burning solution in forest fire prevention.

- With the data collected in February, the model results are the C4 forest fire risk level. At this level, the risk of forest fire is high. Therefore, proactive burning is strictly forbidden. Forest fire prevention and fighting should be strengthened.

- With data collected in March, the model calculated and predicted as a result of the forest fire risk class C5. With forecast class C5, the probability of forest fire is very high.This result is also consistent with the actual situation at this time, under the impact of high temperatures and low humidity, burning materials, the forest can burn at any time.Therefore, it is necessary to take measures to prevent large forest fires, enhance the warnings and prevention of forest fires at the highest level.

When there is a source of fire, the temperature, humidity and the mass of dry material, as well as the burning coefficient K, are the decisive factors for the formation of a fire.However, at the different time, the ability to form a fire is different due to the constant temperature and humidity factors, and the above factors dominate and interact with each other.This also explains that at the same time in different forest plots or in the same forest plot, the probability of forest fire is different. However, at the same time, the difference in fire level of a specific forest area is not large.

Conclusion and Recommendations

Conclusion

The set of variables {***T, H, m1, K***} and the set of variables {***Pc, Tc***} is highly correlated, with the canonical correlation coefficient $\boldsymbol{R} = 0.675$ and the probability significance $\boldsymbol{P} = 3.17 * 10^{-58} << 0.05$. Therefore, it is possible to use the set of variables {T, H, m1, K} in models of forest fire risk forecasting in Bidoup - Nui Ba National Park.

On the basis of discriminant function analysis, three canonical discriminant functions have been established, which are statistically significant ($P << 0.05$). In these 3 functions***CDF1***, ***CDF2*** and ***CDF3***, the Mahalanobis multivariate distance was calculated in the fire risk forecast model in Bidoup - Nui Ba National Park.

According to the results of discriminant function analysis, 5 Fisher's classification functions have been established. Based on these **FCF1**, **FCF2**, **FCF3**, **FCF4** and **FCF5** functions, it is possible to determine the fire risk classes in Bidoup - Nui Ba National Park.

Models of forest fire risk prediction models (Mahalanobis multivariate distance model and Fisher classification function model) in Bidoup - Nui Ba National Park were tested for accuracy and appropriateness from the collected data at different times of the dry season.

Recommendations

These models (Mahalanobis multivariate distance model and Fisher classification function model) can be used to forecast forest fire risk in Bidoup - Nui Ba National Park, serving the develop annual forest fire prevention and fighting plan. These forecasting models can be used in proactive burning solution (prescribed burning). At the time of fires risk class C2 and C3, it is possible to apply the proactive burning solution as the best. Because if burning at the time of the C1 fire forecast class, the layer of fallen leaves can hardly burn well, so it will not reduce the mass of this layer of burning material. In addition, it should be noted that if we proactively burn at the time of C4 and C5 classes, it can cause a very large area of forest fire. This also helps us answer the question in the forest fire research is "when to apply the proactive burning solution in forest fire prevention".

References

1. *Hardle W. &Simar Z.* Applied Multivariate Statistical Analysis. Springer, Berlin, Heidelberg, New York, 2003.
2. *Karouni A., Daya B., Bahlak S.* Forest fire prediction: A comparative study of applicability of fire weather indices for Lebanon // Global Journal on Technology. – 2014. – Vol 5. – P. 8–17.
3. *Mavrakis A., Salvati L.* Analyzing the Behaviour of Selected Risk Indexes During the 2007 Greek Forest Fires // Int. J. Environ. Res. – 2015. – Vol. 9(3). – P. 831–840.
4. *Skvarenina J. Mindas J. Holecy J., Tucek J.* Forest fire weather indices during two largest forest fire events in the Slovak Paradise National Park. – Project VEGA No. 1/9265/02. International Bioclimatological Workshop, Slovakia, 2003.
5. *Le Van Huong* Study on the composition of fire materials of Pinus kesyiaRoyle ex Gordon forest as a basis for proposing fire prevention measures at Bidoup - Nui Ba National Park, Lam Dong province. – Master's thesis in forestry science, Forestry University, 2012 (in Vietnamese)
6. *McLachlan G.J.* Mahalanobis Distance. – Resonance, 1999. – P. 20–26.
7. *Ghosh J.K., Majumdar P.P.* Mahalanobis Prasanta Chandra / In.P. Armitage and T. Colton (Eds), Encyclopedia of Biostatistics. Wiley. NewYork, 1998. – P. 2372–2375.
8. *Mahalanobis P.C.* On tests and measures of group divergence, Journal of the Asiatic Society of Bengal, 1930. – Vol. 26. – P. 541–588.
9. *McLachlan G.J.* Discriminant Analysis and Statistical Pattern Recognition. – New York, Wiley. Technometrics. 1992. – Vol 35 (3). – P. 320 – 330.
10. *Rao C.R.* Linear Statistical Inference and its Applications / Second Edition. – Wiley, New York. – 1973a. – Vol. 57. – Issue 8. – 500 p.
11. *Rao C R, Prasanta Chandra Mahalanobis*. Biographical Memoirs of Fellows of the Royal Society. – 1972. – Vol. 19. – P. 455–492.
12. *Wilks D.S.* Chapter 13 – Canonical Correlation Analysis (CCA) // International Geophysics. – 2011. – Vol 100. – P. 563–582.
13. *Poulsen J, French A*. Discriminant Function Analysis. 2004. – Available at: http://online.sfsu.edu/efc/classes/biol710/discrim/discrim.pdf
14. *Matthew C., Lawoko C.R.O., Korte C.J., Smith D.* Application of canonical discriminant analysis, principal component analysis, and canonical correlation analysis as tools for evaluating differences in pasture botanical composition // New zealand journal of agricultural research. – 1994. – Vol. 37 (4). – P. 509–520.
15. *Edzard V.S.* Application of Canonical Discriminant Analysis for the Assessment of Genetic Variation in Tall Fescue // Crop Science Society of America. Published in Crop Sci. – 2002. – Vol(42). – P. 534–539.
16. *Srivastava M.S.* Methods of Multivariate Statistics. – John Wiley & Sons, Inc., New York, 2002. – 728 p.
17. *Anja Hashagen, Christian Zabel, Heidi Schelhowe, Saeed ZareAdaptive*. Motion Pattern Recognition: Implementing Playful Learning through Embodied InteractionUniversity of Bremen. TZI. dimeb (Digital Media in Education) / Digital Human Modeling: Second International Conference, Held as Part of HCI International 2009 San Diego, CA, USA, Proceedings. Springer, 2009. – P. 104–114.
18. *S. Das Gupta.* Mahalanobis distance / In P Armiage and T Colton (Eds). – Encyclopedia of Biostatistics, Wiley, New York, 1998. – P. 2369–2372.
19. *Cruz-Castillo J.G, Ganeshanandam S., MacKay B.R., Lawes G.S., Lawoko C.R.O., Woolley D.J.* Applications of canonicaldiscriminant analysis in horticultural research // HortScience. – 1994. – Vol 29. – P. 1115–1119.
20. *Dillon W.R., Goldstein M.* Multivariate analysis methods and applications. – John Wiley and Sons, New York. 1984. – Vol. 29 (6). – P. 755–756. https://doi.org/10.1002/bimj4710290617

ПРИМЕНЕНИЕ МНОГОМЕРНЫХ СТАТИСТИЧЕСКИХ МЕТОДОВ ДЛЯ ПРОГНОЗИРОВАНИЯ ЛЕСНОЙ ПОЖАРНОЙ ОПАСНОСТИ В НАЦИОНАЛЬНОМ ПАРКЕ БИДУП-НУЙБА

Ле Ван Хыонг[1], Нгуен Нгок Киенг[2], Нгуен Данг Хой[3], Данг Хунг Куонг[3,4]

[1]Национальный парк Бидуп-Нуйба, деревня Данхим, район Лакдуонг, провинция Ламдонг, Социалистическая Республика Вьетнам

[2]Хошиминский университета сельского и лесного хозяйства, Хошимин, Социалистическая Республика Вьетнам

[3]Совместный Российско-Вьетнамский Тропический научно-исследовательский и технологический центр, г. Ханой, Социалистическая Республика Вьетнам

[4]Южный федеральный университет, Россия, г. Ростов-на-Дону, Российская Федерация

e-mail: lvhuongdl@gmail.com

В статье представлены результаты применения многомерных статистических методов (*CCA*: анализ канонической корреляции и *DFA*: анализ дискриминантной функции) для определения канонической корреляции между набором переменных $\{T, H, m_1, K\}$ и набором переменных $\{Pc, Tc\}$ (T: температура, H: относительная влажность, m_1: масса сухого топлива, K: коэффициент горения, $K = m_1/M$, где M: общая масса горючего топлива, Pc: процент сожженного топлива и Tc: время горения). Были созданы модели прогнозирования лесной пожарной опасности в национальном парке Бидуп-Нуйба по результатам анализа дискриминантной функции DFA. На основании данных исследований за период ноябрь, декабрь, январь, февраль и март 2015–2017 годов из 340 участков для отбора проб (каждый размером 2x2 м) в национальном парке Бидуп-Нуйба мы провели обработку данных в Excel (расчет) и Statgraphics (многомерные статистические методы: *CCA & DFA*). Из нашего анализа были выявлены три результата: *(i)* Каноническая корреляция между набором переменных $\{T, H, m_1, K\}$ и набором переменных $\{Pc, Tc\}$ является очень значимой ($R=0{,}675581$ и $P=3{,}17 * 10^{-58} << 0{,}05$); поэтому мы использовали набор переменных $\{T, H, m_1, K\}$ в моделях прогнозирования опасности лесных пожаров. *(ii)* Были определены коэффициенты стандартизированных и нестандартных канонических дискриминантных функций (*SCDF & UCDF*) и классификационной функции Фишера (*FCF*). *(iii)* Были созданы две модели прогнозирования опасности лесных пожаров (модель расстояния Махаланобиса и модель функции классификации Фишера).

Ключевые слова: Бидуп-Нуйба, канонический корреляционный анализ, анализ дискриминантной функции, классификационная функция Фишера, модели прогнозирования лесной пожарной опасности.

Ле Ван Хыонг	кандидат лесоводских наук, директор Национального парка Бидуп – Нуйба, провинция Лам Донг, e-mail: lvhuongdl@gmail.com
Нгуен Нгок Киенг	доктор сельскохозяйственных наук, Университет сельского и лесного хозяйства г. Хошимин, e-mail: kiengnn@gmail.com
Нгуен Данг Хой	кандидат географических наук, директор Института тропической экологии, Совместный Российско-Вьетнамский Тропический научно-исследовательский и технологический центр, e-mail: danghoi110@gmail.com
Данг Хунг Куонг	магистр географии, научный сотрудник Института тропической экологии, Совместный Российско-Вьетнамский Тропический научно-исследовательский и технологический центр, e-mail: danghungcuong@gmail.com

Поступила в редакцию 02.02.2018 г.