UDC 630*11+57.087.1

A NEW APPROACH TO DEVELOPING A LOGISTIC REGRESSION MODEL VARIABLES TO PREDICT TREE MORTALITY, BASED ON TREE-RING GROWTH DYNAMICS

A. V. Kachaev[1], I. A. Petrov[2], V. I. Kharuk[1, 2], E. N. Belova[1]

[1] *Siberian Federal University*
Svobodny prospekt, 79, Krasnoyarsk, 660041 Russian Federation

[2] *V. N. Sukachev Institute of Forest, Russian Academy of Science, Siberian Branch*
Akademgorodok, 50/28, Krasnoyarsk, 660036 Russian Federation

E-mail: avkachaev@gmail.com, petrovilsoran@gmail.com, v47sugen@gmail.com, enbelova@sfu-kras.ru

The annual tree increment is one of the integral indicators of abiotic and biotic processes occurring in the forest ecosystem. The use of logistic regression models based on annual tree-ring growth data is a promising approach to studying tree mortality. The diversity of logistic variables in scientific research is a result of various choices of statistics (average, median, growth trend, etc.) and their score in the time-window for the past N (5, 10, ..., 40) years. We propose a new scheme for the formation of logistic variables that involves fixing the statistics for calculating the average and choosing two non-intersecting time-windows based on measurements of the annual tree-rings growth. The choice of non-overlapping «windows» enables setting the ratio of the average growth of annual rings of trees between the windows for different periods of time. We examined the past 41 years of tree growth. Logistic regression models are constructed on a set of pairs of non-intersecting «windows» with a limit on the values of the sensitivity and specification of at least 1.6. The calculation of the percentage prediction if a tree is living or dying was done based on the contingency table in the logistic regression model. The logistic regression models were visualized using ROC curves. The models were compared on an expert scale based on the calculated area under the ROC curves. The obtained logistic regression model was verified by the bootstrap method. The calculations were carried out for the Siberian stone pine *Pinus sibirica* du Tour growing in the Baikal region (the Khamar-Daban Ridge) using the R programming language. The computed logistic regression model helped us predict live and dead trees in more than 80 % of cases.

Keywords: *dendrochronology, annual increment, Siberian stone pine Pinus sibirica du Tour, Khamar-Daban.*

How to cite: *Kachaev A. V., Petrov I. A., Kharuk V. I., Belova E. N.* A new approach to developing a logistic regression model variables to predict tree mortality, based on tree-ring growth dynamics // *Sibirskij Lesnoj Zurnal* (Sib. J. For. Sci.). 2020. N. 5. P. 37–44 (in English with Russian abstract and references).

DOI: 10.15372/SJFS20200504

INTRODUCTION

The study of tree mortality is one of the directions of the research of the dynamic processes occurring in the forests ecosystem. In the past two decades, Russia has seen an increase in the area of dead stands as compared with the previous period (Sarnatskiy, 2012; Malakhova, Lyamtsev, 2014; Pavlov, 2015*a*, *b*). The object under study has emergent properties, which causes difficulties in constructing objective models of forest ecosystem (Rozenberg, 2011) and results in a variety of approaches. From the whole variety of research directions, we single out two: constructing mortality models, determined by a mixed set of abiotic and biotic factors (Scott et al., 1999; Greenwood, Weisberg, 2008; Kharuk et al., 2013) and the approach based on dendrochronological measurements (Bigler, Bugmann, 2003; Das et al., 2007; Cailleret et al., 2016). The variables in mixed factors are temperature, precipitation, groundwater level, soil cover patterns, the volume of tree crowns, etc. The dendrochronological approach uses annual tree-ring growth data sets: the width of early, latewood, the ring-width, density, etc.

M. L. Scott et al. (1999) studied the influence of the hydrological regime on the growth of the eastern cottonwood *Populus deltoids* W. Bartram ex Marshall in coastal areas in eastern Colorado, USA. To build a logistic regression model, the following measurements were used: the volume of live crown, the growth of radial stems, and the annual growth of branches for 689 poplar trees. The authors found that a steady decrease in the level of groundwater level, exceeding 1 m, for the poplar tree growing in medium alluvial sands, led to the crown drying and growth of the stems and resulted in up to 88 % of tree mortality over a three-year period.

C. Bigler and H. Bugmann (2003) developed a logistic regression model for the European spruce *Picea abies* (L.) H. Karst. in the subalpine forest of the European Alps. The authors determined the level and growth trend variables in various time windows that are used in the logistic model of tree mortality. Growth levels were calculated as average values of the basal area increment BAI (cm^2/year) for the last 3, 5 or 7 years of the tree (BAI3, BAI5, BAI7). They divided the experimental material into three homogeneous groups in terms of interspecific competition for growth (log BAI3, log BAI5, log BAI7). The slope coefficients for local linear regressions established over the past 5, 10, 15, ... 40 years for BAI (locreg5, ..., locreg40) was defined as growth trends. The developed logistic model based on the level and growth trend variables correctly classified dead or live trees (a plot of 119 trees) in almost 80 % of cases.

A. J. Das et al. (2007) examines mortality of the white fir *Abies concolor* (Gordon) Lindley ex Hildebrand and the sugar pine *Pinus lambertiana* Douglas by developing logistic models using tree growth indices obtained from tree rings over the last N years: average growth (avgN), growth trend (trendN) and a number of sharp drops in growth (abruptN), where N takes values in the interval from 5 to 40 years with a step of 5 years. The developed logistic models correctly classify 78.6 % dead and 83.7 live trees.

In this paper, we developed an approach using dendrochronological measurements. One of the advantages of this approach is that the total number of trees used in the simulation was an order of magnitude smaller than with other approaches to modeling tree mortality.

The relevance of the study is determined by the increase in the percentage of areas with dying trees caused by changes in climatic conditions of the environment (Kharuk et al., 2013, 2017*a*, *b*; Zamolodchikov, Kraev, 2016). The study was aimed at developing a new approach to obtaining the variables (Kachaev, 2019) in logistic regression models of tree mortality using tree rings growth and software implementation in R programming language (Kabakov, 2014).

MATERIALS AND METHODS

We chose to develop the logistic regression model for the Siberian stone pine *Pinus sibirica* du Tour growing in the Khamar-Daban mountain ridge (Kharuk et al., 2017*b*).

In the article (Kharuk et al., 2017*b*) the causes of die-off of the Baikal cedar forests (the Khamar-Daban mountain ridge) were examined by satellite remote sensing and the dendrochronological method. Since the 1980s a decrease in the value of the growth index (R^2 = 0.69) and a decrease of the dryness climate index SPEI (R^2 = 0.72) have been established. In the mid-2000s an increase in aridity led to the division of cedar trees into two groups: «survivors» and «drying». The spatial distribution of these groups is associated with orographic factors (exposure, slope, height above sea level, etc.), which lead to a moisture deficit. The tree growth index is closely related to the June dryness index (R^2 = 0.55). Along with water stress, drying trees are exposed to stem pests and worms and pathogens. It has been established that the primary cause of cedar pine drying is water stress due to an increase in aridity of the climate. In general, within the Khamar-Daban Ridge, heavily damaged and dying trees (> 50 % of dying and dead trees) make up 8–10 % of the total area of dark coniferous.

The characteristics of the materials used are summarized in Table 1.

The number of living trees is 17 % less than that of dead trees.

Live and dead trees sampled had similar dynamics regarding age and radial increment, as is clear from Fig. 1 and 2.

Table 1. The main characteristics of the source data. Specie – Siberian stone pine

Site	Dead	Live	DBH	Age, years
Khamar-Daban Ridge	83	–	18.68–57.15	60–247
	–	69	13.46–59.61	54–263

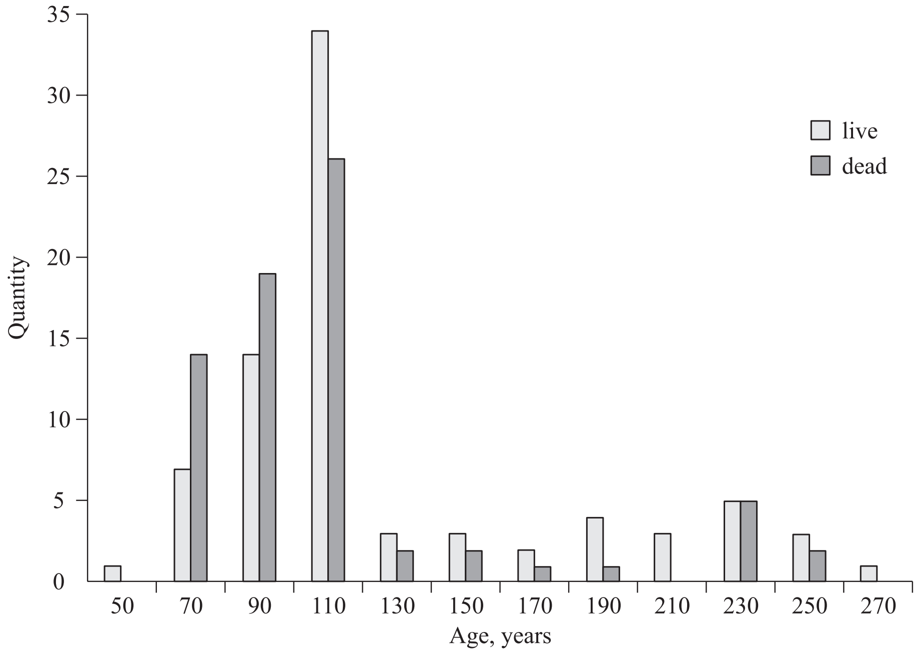

Fig. 1. The distribution of the number of live and dead trees in the age groups of 20 years.

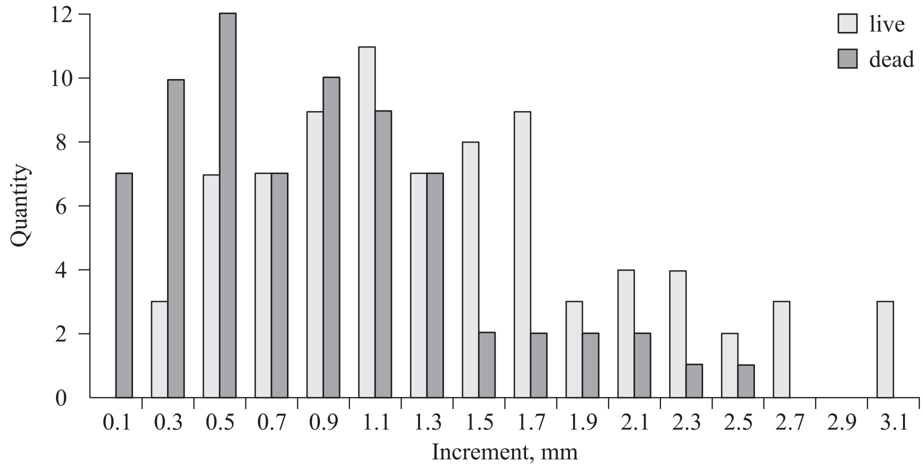

Fig. 2. Distributions of the average increment over the past 5 years.

Overlapping radial increments of living and dead trees did not allow us to divide all the trees by the average increment into two groups for the past few years (for example, 5 years), as can be seen from Fig. 3.

Description of method. To develop logistic regression models, we introduce the following sets and notation.

Suppose $Gr_{i,p} = \{g_{ri,p}, gr_{i,p-1}, \ldots, gr_{i,1}\}$ is a dendrochronological series of growth with the reverse indexing, where p – is the length of the series, i – the number of the series. By $Gr = \{Gr_{i,p} \mid i \in [1, n]\}$ we label a set of dendrochronological series of growth, where $n = \overline{\overline{Cr}}$ is the quantity of series.

The «status» $St = \{s_p, s_{p-1}, \ldots, s_1 \mid s_j = 0$ или $1, j \in [1, p]\}$ is connected with each dendrochronological series. «Status» 1 means a live tree and «status» 0 – a dead tree.

We define «window» $W_{c,w} = \{c + [w/2], \ldots, c + 1, c, c - 1, \ldots, c - [w/2]\}$ as a number sequence, where c – is the center of the window, w – the window size and $[w/2]$ – a quotient of fraction $w/2$. A «window» is a vector of numbers that denote index numbers of a dendrochronological series. $W^p = \{W_{c,w}\}$ labels a set of «window» for series of length p. Suppose $m = \overline{\overline{W^p}}$ is a number of «windows». For further calculations we introduce the following limitations for all time series: the series length is $p = 41$,

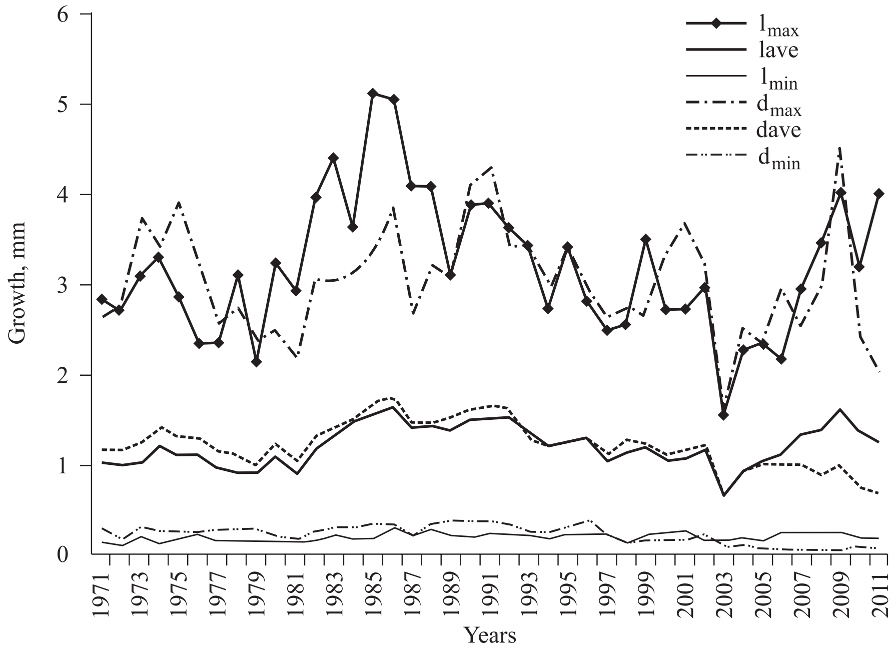

Fig. 3. The distribution of the average increment over the sliding «window» of 5 years in the group of live (lave) and dying (dave) trees by year lmax and lmin is the maximum and minimum increase in the group of living trees and d_{max} and d_{min} is the maximum and minimum increase in the group of drying trees.

the window size w is an odd ordinal from 5 to 21. With these restrictions, the number of «windows» $m = 361$.

Then we define the mean of the window $Mean_{i,W_{c,w}} = \frac{\Sigma j \in W_{c,w} gr_{i,j}}{w}$, where i – is the number of series. $Mean_{W_{c,w}} = \{Mean_{i,W_{c,w}} \mid i \in [1, n]\}$ labels a set of window statistics vectors for all series. It is clear that the number of window statistics vectors equals the number of «window», i. e. 361. Further, the vector of window statistics determines a variable in the logistic regression model. Note that the total number of logistic regression models with two variables is equal to the number of combinations $C^2_{361} = 64\,980$.

In our study, we restrict the search for logistic regression models for «window» without intersection points. The number of all pairs of non-intersecting windows is programmatically computed and makes 14 688.

Logistic regression is a statistical model that is used to predict the probability of an event occurrence by the values of a set of features:

$$P(Y) = \frac{1}{1+e^{-f(X)}}, \quad (1)$$

where $P(Y|X)$ – a probability of the event with the feature vector $X = (X_1, X_2 \ldots X_n)$ equals 1 or 0, $f(X) = b_0 + \sum_{i=1}^{n} b_i X_i$ – is a linear function from X_i features.

Calculating coefficients of logistic regression, a set of features X is defined by a set of window statistics $Mean_{W_{c,w}}$.

To visualize the results of logistic regression was used ROC – a curve, often used to evaluate the quality of binary classification (Davis, Goadrich, 2006).

As a result of logistic regression classification of tree drying, on the basis of obtained measurements of tree growth, we can conclude whether the tree is live or dead. Four variants are possible:

– *TP* (true-positive) – tree is dying, positive conclusion,

– *FP* (false-positive) – tree is live, positive conclusion,

– *TN* (true-negative) – tree is live, negative conclusion,

– *NP* (false-negative) – tree is dying, negative conclusion.

The above conclusions are presented in the form of a contingency table (Table 2).

Table 2. Contingency table

Classification	Real situation	
	positive	negative
Positive	*TP*	*FP*
Negative	*FN*	*TN*

During the analysis comparative rather than invariable ratios expressed in percentage are used. The fraction of true positive conclusions is defined as $TPR = \dfrac{TP}{TP + FN} \cdot 100\,\%$, the fraction of false positive conclusions is $FPR = \dfrac{FP}{TP + FP} \cdot 100\,\%$.

ROC is a curve presenting the ratio between the part of objects from the total number of carriers of the trait, truly classified as *TPR* (true positive rate) – carriers, called *sensitivity* of classification algorithm, and the part of objects from the total number of non-carriers of the trait, falsely classified as *FPR* (false positive rate) – carriers, the value 1-*FPR* is defined as *specificity* of classification algorithm with variations of decision rule threshold.

The area under ROC-curve called AUC (Area Under Curve) allows numerical comparison of different models. In literature (Hosmer, Lemeshow, 2000) the following expert scale for AUC values is presented that helps evaluate the quality of the model:

– if AUC = 0.5 there is no partitioning (discrimination), it is like tossing a coin,

– if 0.7 ≤ AUC < 0.8 partitioning is acceptable,

– if 0.8 ≤ AUC < 0.9 partitioning is excellent,

– if AUC ≥ 0.9 partitioning is ideal.

To compare the models AIC – Akaike information criterion (Mastitskiy, Shitikov, 2014) is used: the less the value the better. The absolute value of the criterion has no sense.

RESULTS AND DISCUSSION

Initial data of live and dead trees are stored in two .rwl files (Speer, 2010) HDcdead.rwl and HDcliving.rwl respectively. For the convenience of program processing we made a .json file (Vvedeniye v JSON, 2019) – HDc.json, containing the information on the status («live», «dead») of the trees and growth series of all the trees.

In the program chain of search for logistic regressions, we select the following program blocks:

– formation of a system of non-overlapping windows,

– search for logistic regression modelsglm (Mastitskiy, Shitikov, 2014) with condition for an AUC value or sens + spec sum,

– verification of the selected logistic regression models with the bootstrap method (Mastitskiy, Shitikov, 2014),

– visualization of the estimated data.

Below in Table 3 we present the search for models with one variable of all 361 window variants.

Table 3. The results of search for logistic models with one variable

N	Formula	Sums	Sens	Spec	AUC	AIC
1	ttre ~ $W_{3.5}$	1.33	0.68	0.65	0.73	190.28
2	ttre ~ $W_{4.5}$	1.26	0.64	0.63	0.69	197.50
3	ttre ~ $W_{4.7}$	1.29	0.66	0.63	0.69	197.87

Table 4. Formulae of logistic regressions and statistics

N	W_1		W_2		Sens	Spec	Sum	AUC	AIC
	c	*w*	*c*	*w*					
1	3	5	10	9	0.850	0.764	1.614	0.885	144.241
2	3	5	9	7	0.825	0.806	1.631	0.882	144.645
3	3	5	11	11	0.825	0.792	1.617	0.881	146.245

Note. * W_1, W_2 – the first and second windows; *c* – the center of the window; *w* – window size; sums = sens + spec; sens – sensitivity; spec – specificity.

This is the method of full enumeration with AUC values restricted to greater than or equal to 0.69. Note, in (Das et al., 2007) only 8 window sizes are stated (5, 10, 15, 20, 25, 30, 35, 40).

Table 4 presents the result of search for logistic models with two variables by the method of full enumeration with AUC values restricted to greater than or equal to 0.881.

Fig. 4 presents the shaded areas for windows $W_{3.5}$ and $W_{10.9}$ for variant 1 from Table 4, the solid line for live trees, the dotted line for dead trees.

Now we build a line of division for logistic regression with variables: ttre ~ Mean$W_{3.5}$ + Mean$W_{10.9}$.

The logistic regression equation for the line Fig. 5 has the following form:

$$Y = 0.121 + 5.36 \cdot X_1 - 5.88 \cdot X_2, \qquad (2)$$

where variable X_1 – is the mean window $W_{3.5}$, X_2 – is the mean window $W_{10.9}$.

The percentage of truly positive conclusions of live trees makes *TPR* = 82.09 %, and falsely positive *FPR* = 23.61 %. The percentage of truly positive conclusions of dead (drying) trees makes *TNR* = 85 %, and falsely positive *FNR* = 15 %.

Confidence error for the average value of AUC is 0.888 ± 0.001.

«Currently, it is not possible to carry out a meta-analysis of the growth-mortality relationships already published, as most of the authors followed different methodologies to derive them» (Cailleret et al., 2016).

M. Cailleret et al. (2016) distinguish the differences both in the choice of variables, schemes

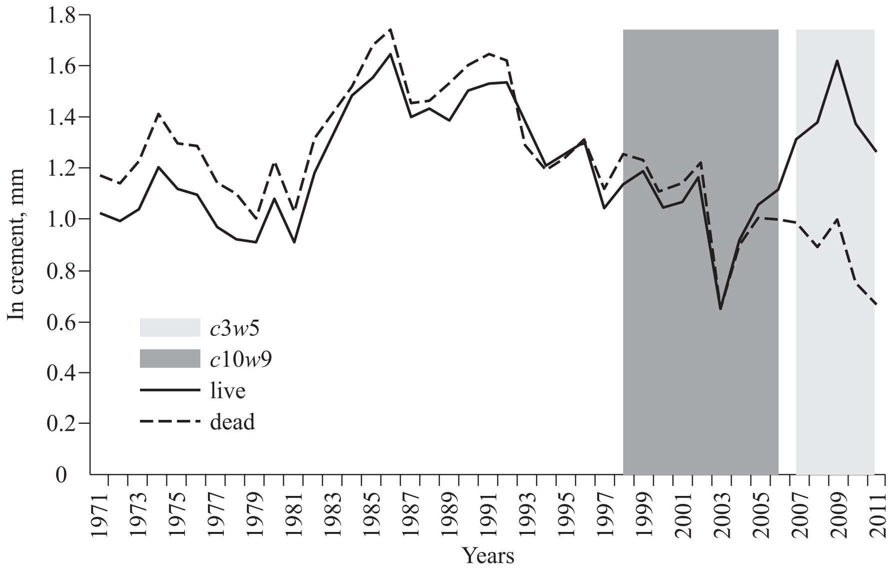

Fig. 4. The average increment for the group of live and dead (drying) trees.

of logistic regression models and in the volume of basic data from tens to thousands of live and dead (drying) trees. Further, we classify the variables analyzed in M. Cailleret et al. (2016).

As basic data we take dendrochronological series: the ring-width (RW) or the basal area increment (BAI) of trees. The diversity of variables in the logistic models resulted from the values of the time-window functions of RW or BAI series. The variety of the logistic model variables was determined by the values of «window» statistics from the RW or BAI series. We took the formula for the

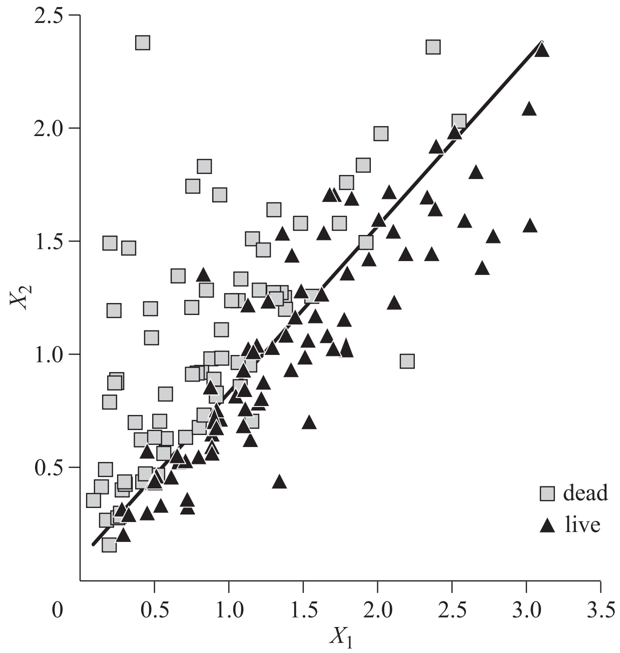

Fig. 5. The line of the division of the logistic regression between live and dead (drying) trees with average increment in windows $W_{3.5}$ and $W_{10.9}$.

variables of the «window» statistics from M. Cailleret et al. (2016) and designated them as statRowN, where stat is the statistic, Row takes the value RW or BAI, and N is the window size. The statistics are functions: mean (average), slo (the linear regression slope), SD (standard deviation), median (median) and others. Examples of variables of «window» statistics: sloRW25, meanBAI35, medianRW17, etc. Thus, in the article (Das et al., 2007) models are defined by the variables: avgN (the mean in the window for the last N years), trendN (the age trend for the last N years) and abruptN (the number of decreases for the last N years), where N is the time interval from 5 to 40 in increment of 5 years.

The originality of our approach is in the fixed statistics and in the choice of two different time-windows for calculating variables for logistic regression modelling. As statistics the mean window was taken. Using Occam's razor, stating «Do not multiply entities beyond what is necessary», we calculate logistic regressions with two values of average growth in different windows for each tree.

We compare our logistic models with one and two variables with the results described in A. J. Das et al. (2007). The basic data in A. J. Das et al. (2007) is consistent with ours: 77 live and 78 dead trees and 80 live and 72 dead trees in this study. The model with one variable avg5, the value of AUC = 0.747 (Das et al., 2007) is well compared with our model with one variable Mean$W_{3.5}$ and the value of AUC equal to 0.73. Note that the variables avg5 and Mean$W_{3.5}$ label the mean window for the last 5 years.

Our logistic models with two variables have a larger AUC value than the model of A. J. Das et al.

(2007), 0.888 vs 0.800, that used three variables, avg10, trend40 and abrupt5.

Obviously, our solution has an advantage over the approach of A. J. Das et al. (2007) and M. Cailleret et al. (2016) in that we use only one statistic of the average with visualization of a two-dimensional representation of the logistic model, whereas they used three and more statistics to result in a more complicated model interpretation process.

CONCLUSION

The novelty of the suggested method is provided by the choice of two different (non-overlapping) windows and one type of statistical data (the mean) for building of variables for logistic regression using tree growth series.

REFERENCES

Bigler C., Bugmann H. Growth-dependent tree mortality models based on tree rings // Can. J. For. Res. 2003. V. 33. N. 2. P. 210–221.

Cailleret M., Bigler C., Bugmann H., Camarero J. J., Cúfar K., Davi H., Mészáros I., Minunno F., Robert E. M., Suarez M. L., Tognetti R., Martínez-Vilalta J. Towards a common methodology for developing logistic tree mortality models based on ring-width data // Ecol. Appl. 2016. V. 26. Iss. 6. P. 1827–1841.

Das A. J., Battles J. J., Stephenson N. L., Mantgem P. J. van. The relationship between tree growth patterns and likelihood of mortality: a study of two tree species in the Sierra Nevada // Can. J. For. Res. 2007. V. 37. N. 3. P. 580–597.

Davis J., Goadrich M. The relationship between precision-recall and ROC curves // Proc. 23rd Int. Conf. Machine Learning. 2006. P. 233–240.

Greenwood D. L., Weisberg P. J. Density-dependent tree mortality in pinyon-juniper woodlands // For. Ecol. Manag. 2008. V. 255. N. 7. P. 2129–2137.

Hosmer D. W., Lemeshow S. Applied logistic regression. 2 ed. New York: J. Wiley & Sons, Inc., 2000. 375 p.

Kabakov R. I. R v deystvii. Analiz i vizualizatsiya dannykh v programme R / per. s angl. P. A. Volkovoy (R in action. Analysis and visualization of data in the program R / transl. from English P. A. Volkova). Moscow: DMK Press, 2014. 588 p. (in Russian).

Kachaev A. V. O vybore peremennykh v logisticheskikh regressionnykh modelyakh usykhaniya derevyev (On the choice of variables in logistic regression models of mortality of trees) // Lesnye ekosistemy borealnoy zony: bioraznoobrazie, bioekonomika, ekologicheskie riski: mat-ly Vseros. konf. s mezhdunar. uchast. (Forest ecosystems of the boreal zone: biodiversity, bioeconomics, environmental risks. Proc. All-Rus. Conf. Int. Participat.). 2019. P. 165–168 (in Russian with English abstract).

Kharuk V. I., Im S. T., Oskorbin P. A., Petrov I. A., Ranson K. J. Siberian pine decline and mortality in southern Siberian Mountains // For. Ecol. Manag. 2013. V. 310. P. 312–320.

Kharuk V. I., Im S. T., Petrov I. A., Dvinskaya M. L., Fedotova E. V., Ranson K. J. Fir decline and mortality in the southern Siberian mountains // Reg. Environ. Change. 2017*a*. V. 17. N. 3. P. 803–812.

Kharuk V. I., Im S. T., Petrov I. A., Golyukov A. S., Ranson K. J., Yagunov M. N. Climate-induced mortality of Siberian pine and fir in the Lake Baikal watershed, Siberia // For. Ecol. Manag. 2017*b*. V. 384. P. 191–199.

Malakhova E. G., Lyamtsev N. I. Rasprostranenie i struktura ochagov usykhaniya elovykh lesov Podmoskovya 2010–2012 godakh (Extent and structure of Moscow region spruce forest dieback in 2010–2012) // Izv. SPb. Lesotekh. akad. (Bull. St. Petersburg For. Acad.). 2014. Iss. 207. P. 193–201 (in Russian with English abstract).

Mastitskiy S. E., Shitikov V. K. Statistichesky analiz i vizualizatsiya dannykh s pomoshchyu R (Statistical analysis and data visualization using R). Moscow, 2014. 401 p. (in Russian).

Pavlov I. N. Bioticheskie i abioticheskie faktory usykhaniya khvoynykh lesov Sibiri i Dalnego Vostoka (Biotic and abiotic factors as causes of coniferous forests dieback in Siberia and Far East) // Sib. ekol. zhurn. (Sib. J. Ecol.). 2015*a*. V. 22. N. 4. P. 537–554 (in Russian with English abstract).

Pavlov I. N. Biotic and abiotic factors as causes of coniferous forests dieback in Siberia and Far East // Contemp. Probl. Ecol. 2015*b*. V. 8. N. 4. P. 440–456 (Original Rus. text © I. N. Pavlov, 2015, publ. in Sibirskij ekologicheskij zhurnal. 2015*a*. V. 22. N. 4. P. 537–554).

Rozenberg G. S. Ekologiya i fizika: paralleli ili seti? (v prodolzhenie diskussii) (Ecology and physics: parallels or network? (pending discussion)) // Biosfera (Biosphere). 2011. V. 3. N. 3. P. 296–303 (in Russian with English abstract).

Sarnatskiy V. V. Zonal'no-tipologicheskie zakonomernosti periodicheskogo massovogo usykhaniya el'nikov Belarusii (Zonal-typological patterns of periodic mass drying of spruce forests of Belarus) // Tr. BGTU (Proc. Belarus St. Technol. Univ.). N. 1. Lesn. khoz-vo (Forestry). 2012. V. 148. N. 1. P. 274–276 (in Russian with English abstract).

Scott M. L., Shafroth P. B., Auble G. T. Responses of riparian cottonwoods to alluvial water table declines // Environ. Manag. 1999. V. 23. P. 347–358.

Speer J. H. Fundamentals of tree-ring research. Tucson, AZ: Univ. Arizona Press, 2010. 368 p.

Vvedeniye v JSON (Introduction to JSON), 2019 (in Russian). https://www.json.org/json-ru.html

Zamolodchikov D., Kraev G. Vliyanie izmeneniy klimata na lesa Rossii: zafiksirovannye vozdeystviya i prognoznye otsenki (Influence of climate change on Russian forests: recorded impacts and forecast estimates) // Ustoychivoe lesopol'zovanie (Sustainable forest management). 2016. N. 4 (48). P. 23–31 (in Russian).

УДК 630*11+57.087.1

НОВЫЙ ПОДХОД К ФОРМИРОВАНИЮ ПЕРЕМЕННЫХ ЛОГИСТИЧЕСКОЙ РЕГРЕССИОННОЙ МОДЕЛИ ПРОГНОЗА УСЫХАНИЯ ДЕРЕВЬЕВ НА ОСНОВЕ ДИНАМИКИ ГОДИЧНЫХ КОЛЕЦ

А. В. Качаев[1], И. А. Петров[2], В. И. Харук[1, 2], Е. Н. Белова[1]

[1] *Сибирский федеральный университет*
660041, Красноярск, просп. Свободный, 79

[2] *Институт леса им. В. Н. Сукачева СО РАН – обособленное подразделение ФИЦ КНЦ СО РАН*
660036, Красноярск, Академгородок 50/28

E-mail: avkachaev@gmail.com, petrovilsoran@gmail.com, v47sugen@gmail.com, enbelova@sfu-kras.ru

Поступила в редакцию 20.07.2020 г.

Годичный прирост деревьев является одним из интегральных показателей абиотических и биотических процессов, происходящих в лесной экосистеме. Один из подходов к изучению процессов смертности деревьев – построение логистических регрессионных моделей с использованием годичного прироста. Многообразие формирования логистических переменных в научных исследованиях определяется различным выбором статистик (среднее, медиана, тренд роста и т. д.) и счетом их в «окне» за последние 5, 10, …, 40 лет. Нами предлагается новая схема формирования логистических переменных, фиксирующая статистику вычисления среднего и выбора двух непересекающихся «окон» по измерениям приростов годичных колец. Выбор непересекающихся «окон» позволяет задать отношение среднего прироста годичных колец деревьев в разные периоды. Нами исследуется последний период (41 год) прироста деревьев. На множестве пар непересекающихся «окон» полным перебором строятся логистические регрессионные модели с ограничением на значения суммы чувствительности и спецификации не менее 1.6. Расчет процента предсказания «живое» дерево или «усыхает» определяется через таблицу сопряженности в логистической регрессионной модели. Визуализация логистических регрессионных моделей осуществлена с использованием ROC-кривых. Модели сравниваются по экспертной шкале на основании рассчитанной площади под ROC-кривыми. Верификация логистической регрессионной модели проведена с использованием бутстрэп-метода. Расчеты выполнены с использованием языка программирования *R* для деревьев сосны кедровой сибирской *Pinus sibirica* du Tour, произрастающих в Прибайкалье (хр. Хамар-Дабан). Построенная логистическая регрессионная модель предсказывает живые и усыхающие деревья в более чем в 80 % случаев.

Ключевые слова: *дендрохронология, годичный прирост, сосна кедровая сибирская Pinus sibirica du Tour, Хамар-Дабан.*