SYLWAN 157 (9): 652-661, 2013

RAFAŁ PODLASKI, FRANCIS A. ROESCH

Aproksymacja rozkładów pierśnic drzew w drzewostanach dwugeneracyjnych za pomocą rozkładów mieszanych. II. Testy zgodności

Approximation of the breast height diameter distribution of two-cohort stands by mixture models. II. Goodness-of-fit tests

ABSTRACT

Podlaski R., Roesch F. A. 2013. Aproksymacja rozkładów pierśnic drzew w drzewostanach dwugeneracyjnych za pomocą rozkładów mieszanych. II. Testy zgodności. Sylwan 157 (9): 652-661.

The goals of this study are (1) to analyse the accuracy of the approximation of empirical distributions of diameter at breast height (dbh) using two-component mixtures of either the Weibull distribution or the gamma distribution in two-cohort stands, and (2) to discuss the procedure of choosing goodness-of-fit tests. The study plots were located in the Świętokrzyski National Park (central Poland) and in the Southern Appalachian Mountains (eastern USA). The results of the goodness-of-fit tests (chi-squared, Kolmogorov-Smirnov, Cramér-von Mises, and Anderson-Darling), normalised bias and normalised root mean square error, indicate that dbh empirical distributions of two-cohort stands are compatible with the mixture models investigated. The chi-squared test and the generalization of the Anderson-Darling test to discrete distributions should be used to assess whether empirical dbh data are consistent with a hypothesized null distribution.

KEY WORDS

two-component mixtures, tree diameter distribution, nonparametric goodness-of-fit tests, discrete null distribution

ADDRESSES

Rafał Podlaski [1] – e-mail: r_podlaski@pro.onet.pl
Francis A. Roesch [2] – e-mail: froesch@fs.fed.us

[1] Zakład Ochrony Przyrody; Uniwersytet Jana Kochanowskiego; ul. Świętokrzyska 15; 25-406 Kielce

[2] Southern Research Station, USDA Forest Service; 200 WT Weaver Blvd., Asheville, NC 28804-3454 USA

Wstęp

Analiza rozkładów pierśnic jest wykorzystywana w teorii i praktyce leśnej m.in. do oceny budowy drzewostanów oraz do modelowania wzrostu drzew [Bruchwald 2001; Bruchwald, Zasada 2010; Bruchwald i in. 2011]. Modele mieszane, w porównaniu do rozkładów pojedynczych, są znacznie bardziej uniwersalne i pozwalają na istotne zwiększenie stopnia dopasowania w przypadku aproksymacji rozkładów dwumodalnych [Jaworski, Podlaski 2012]. Złożone z dwóch rozkładów składowych rozkłady mieszane okazały się bardzo użyteczne zarówno w trakcie aproksymacji danych empirycznych [Liu i in. 2001; Zhang i in. 2001; Gove i in. 2008; Podlaski 2010a, b, 2011a, b], jak i podczas wydzielania podpopulacji związanych z generacjami wiekowymi i/lub z poszczególnymi gatunkami drzew [Zasada, Cieszewski 2005].

Testy zgodności weryfikują hipotezy dotyczące rozkładu zmiennej losowej. Celem tych testów jest m.in. porównanie rozkładu empirycznego z rozkładem teoretycznym (testy dla jednej próby). Ogólna ich idea bazuje na założeniu, że jeżeli dana cecha empiryczna ma określony rozkład teoretyczny, to wartości pewnych statystyk obliczonych dla danych empirycznych i dla rozkładu teoretycznego powinny się niewiele różnić. W statystyce stosowane są różne testy zgodności różniące się m.in. założeniami co do charakteru danych empirycznych (dane ciągłe czy skokowe) oraz mocą [D'Agostino, Stephens 1986; Law, Kelton 2000]. Testowanie zgodności empirycznych rozkładów pierśnic z teoretycznymi rozkładami mieszanymi wymaga na ogół stosowania testów uniwersalnych, pozwalających na obliczenie granicznego poziomu istotności (wartości p) dla dowolnych rozkładów [de Wet, Venter 1994].

Podczas oceny stopnia dopasowania rozkładu teoretycznego do rozkładu empirycznego, oprócz testów zgodności, stosowane są również inne statystyki. Najczęściej wykorzystywane są: (1) indeksy charakteryzujące sumę różnic między empiryczną i teoretyczną liczbą pierśnic, wielkością pola powierzchni przekroju pierśnicowego czy miąższością w poszczególnych stopniach grubości (np. obciążenie, błąd średniokwadratowy, indeks błędu Reynoldsa), (2) statystyki bazujące na wartościach funkcji wiarygodności osiągającej ekstremum globalne, (3) statystyki uwzględniające współczynnik korelacji między prawdopodobieństwem empirycznym i teoretycznym i (4) metody graficzne prezentujące płaszczyzny utworzone dla współczynnika skośności i spłaszczenia [Reynolds i in. 1988; Li i in. 2002; Wang, Rennolls 2005].

Celem pracy jest analiza dokładności aproksymacji rozkładów empirycznych pierśnic rozkładami mieszanymi, złożonymi z dwóch rozkładów składowych Weibulla i dwóch rozkładów składowych gamma, w drzewostanach dwugeneracyjnych oraz przedyskutowanie sposobów postępowania podczas wyboru testów zgodności dla jednej próby, pozwalających na ocenę dokładności aproksymacji. Drzewostany dwugeneracyjne odznaczają się na ogół dwupiętrową budową i zróżnicowaną teksturą [Korpeľ 1995; Olano, Palmer 2003; Jaworski, Podlaski 2007].

Materiał i metody

Powierzchnie badawcze zostały założone w Górach Świętokrzyskich (50°50'-50°53'N, 20°48'-21°05'E; 11 powierzchni) oraz w Appalachach Południowych (34°59'-36°32'N, 78°43'-84°13'W; 19 powierzchni). Na każdej z nich zmierzono pierśnice wszystkich żywych drzew. Szczegółowy opis zastosowanej metodyki znajduje się w pracy Podlaskiego i Roescha [2013]. W Górach Świętokrzyskich liczba drzew na powierzchniach badawczych wynosiła od 93 (powierzchnia ŚPN24) do 188 (powierzchnia ŚPN21). Średnia liczba drzew na powierzchni to 126,6 drzewa. W Appalachach Południowych liczba drzew na powierzchniach badawczych kształtowała się od 21 (powierzchnia 39065) do 48 (powierzchnia 193028), a średnia wynosiła 32 drzewa.

Badane drzewostany różnią się istotnie m.in. warunkami topograficznymi i siedliskowymi. W Górach Świętokrzyskich najwyższym wzniesieniem jest Łysica mająca 612 m n.p.m., a w Appalachach Południowych obszar objęty badaniami wykazuje zróżnicowanie wysokości od ok. 300 do 2040 m n.p.m. Podobnie średnia temperatura stycznia i lipca wynosiła w Górach Świętokrzyskich odpowiednio –5 i 16°C (dane ze stacji na Świętym Krzyżu), natomiast w Appalachach Południowych przyjmowała wartości od 2 do 10°C (w styczniu) oraz od 18 do 31°C (w lipcu). W Górach Świętokrzyskich rośnie 35 rodzimych gatunków drzew, podczas gdy w Appalachach Południowych występuje 100 gatunków [Gądek 2000; http://www.nps.gov].

Do modelowania empirycznych rozkładów pierśnic zastosowano rozkłady mieszane złożone z dwóch rozkładów składowych Weibulla i dwóch rozkładów składowych gamma:

$$f_{(Weib)X}(x|\psi) = \pi_1 f_{(Weib)1}(x|\theta_1) + \pi_2 f_{(Weib)2}(x|\theta_2) \qquad [1]$$

$$f_{(gam)X}(x|\psi) = \pi_1 f_{(gam)1}(x|\theta_1) + \pi_2 f_{(gam)2}(x|\theta_2) \qquad [2]$$

gdzie:

π_1 i π_2 to odpowiednio wagi (frakcje) rozkładów składowych $f_{(\cdot)1}(\cdot)$ i $f_{(\cdot)2}(\cdot)$, θ_1 i θ_2 to wektory parametrów rozkładów składowych, ψ to zestaw wszystkich parametrów danego rozkładu mieszanego.

Funkcje gęstości pojedynczego rozkładu Weibulla i gamma mają odpowiednio następującą postać:

$$f_{(Weib)X}(x\,|\,\alpha,\beta,\gamma)=\frac{\alpha}{\beta}\left(\frac{x-\gamma}{\beta}\right)^{\alpha-1}e^{-\left(\frac{x-\gamma}{\beta}\right)^{\alpha}} \quad [3]$$

$$f_{(gam)X}(x\,|\,\alpha,\beta,\gamma)=\frac{\beta^{\alpha}(x-\gamma)^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta(x-\gamma)} \quad [4]$$

gdzie:

parametr kształtu α>0,
parametr skalujący β>0,
$x \geq$ od parametru przesunięcia γ,
$\Gamma(\cdot)$ – funkcja gamma.

Do oceny zgodności dopasowania rozkładów teoretycznych do rozkładów empirycznych wykorzystano test chi-kwadrat [Macdonald, Pitcher 1979; Reynolds i in. 1988]:

$$\chi^2=-2\sum_{j=1}^{l}n_j\log\left(\frac{\hat{n}_j}{n_j}\right) \quad [5]$$

gdzie:

$\hat{n}_j$ i n_j – oznaczają odpowiednio oszacowaną (na podstawie rozkładu teoretycznego) i rzeczywistą liczbę drzew w j-tym stopniu grubości,
l – to liczba stopni grubości.

Test chi-kwadrat ma k=l-np-1 stopni swobody (np – liczba estymowanych parametrów).

Ponadto do oceny zgodności zastosowano trzy testy bazujące na analizie różnic między dystrybuantą empiryczną $F_{dbh}(x)$ i teoretyczną $F_{T(\cdot)}(x)$ [Cramér 1928; von Mises 1928; Anderson, Darling 1952; Slakter 1965]:

– Kołmogorowa-Smirnowa:

$$D=\sup_x\left|F_{dbh}(x)-F_{T(\bullet)}(x)\right| \quad [6]$$

– Craméra-von Misesa:

$$W^2=n\int_{-\infty}^{+\infty}\left[F_{dbh}(x)-F_{T(\bullet)}(x)\right]^2 dF_{T(\bullet)}(x) \quad [7]$$

– Andersona-Darlinga:

$$A^2=n\int_{-\infty}^{+\infty}\left[\frac{F_{dbh}(x)-F_{T(\bullet)}(x)}{F_{T(\bullet)}(x)-F_{T(\bullet)}(x)^2}\right]^2 dF_{T(\bullet)}(x) \quad [8]$$

gdzie:

n – liczba zmierzonych pierśnic.

Dla testu chi-kwadrat nie musi być spełniony warunek ciągłości dystrybuanty, natomiast pozostałe trzy testy zgodności wymagają ciągłych dystrybuant i dlatego zastosowano ich zmodyfikowane wersje, możliwe do wykorzystania również dla danych skokowych. Szczegółowy opis wersji testu Kołmogorowa-Smirnowa dla cech skokowych znajduje się w pracach Conovera [1972] i Glesera [1985], natomiast wersja testu Craméra-von Misesa i Andersona-Darlinga, umożliwiająca analizę cech skokowych, bazuje na pracy Choulakiana i in. [1994].

Obliczenia przeprowadzono wykorzystując pakiety *mixdist* [Macdonald, Du 2004] oraz *dgof* [Arnold, Emerson 2011], wchodzące w skład środowiska R (http://www.Rproject.org). W celu określenia obciążenia i błędu aproksymacji wykorzystano znormalizowane obciążenie (NB) i znormalizowany pierwiastek z błędu średniokwadratowego (NRMSE), określone wzorami:

$$\mathrm{NB} = \frac{1}{(d_{c\max} - d_{c\min})\, l} \sum_{j=1}^{l} (n_j - \hat{n}_j) \qquad [9]$$

$$\mathrm{NRMSE} = \frac{1}{(d_{c\max} - d_{c\min})} \sqrt{\frac{1}{l} \sum_{j=1}^{l} (n_j - \hat{n}_j)^2} \qquad [10]$$

gdzie:

d_{cmin} i d_{cmax} to odpowiednio środek najmniejszej i największej klasy grubości, a pozostałe oznaczenia jak w przypadku testu chi-kwadrat.

Wyniki

Wartości statystyki χ^2 i graniczny poziom istotności (p>0,05) świadczą o poprawnym aproksymowaniu danych empirycznych przez rozkład mieszany Weibulla w przypadku 22 powierzchni na 29, dla których obliczono parametry tego rozkładu, i przez rozkład mieszany gamma w przypadku 24 na 30 powierzchni (tab. 1, 2). Wartości statystyk D, W^2 i A^2 oraz odpowiednich granicznych poziomów istotności dowodzą precyzyjnego wyrównania danych empirycznych przez analizowane rozkłady mieszane w przypadku wszystkich badanych drzewostanów (tab. 1, 2). Zastosowane testy zgodności wykazały, że najmniej dokładnie rozkłady empiryczne zostały aproksymowane modelami teoretycznymi dla powierzchni ŚPN25, ŚPN23, 75045, ŚPN31, 23046 (w przypadku rozkładu mieszanego Weibulla) oraz dla powierzchni 75045, ŚPN23, ŚPN28, ŚPN25, 23046, ŚPN31 (w przypadku rozkładu mieszanego gamma). Test chi-kwadrat wskazywał dodatkowo na duże różnice między odpowiednimi rozkładami empirycznymi i teoretycznymi dla powierzchni ŚPN28, 27013 i ŚPN27 (rozkład mieszany Weibulla) oraz dla powierzchni 27013 i ŚPN27 (rozkład mieszany gamma). Najlepsze dopasowanie dla obu uwzględnionych rozkładów mieszanych uzyskano dla powierzchni 75028 (tab. 1, 2). Dokładniejsze aproksymacje wykazano dla powierzchni z Appalachów Południowych w porównaniu do powierzchni z Gór Świętokrzyskich oraz dla rozkładu mieszanego Weibulla w porównaniu do rozkładu mieszanego gamma (różnice były niewielkie).

Wyniki otrzymane z wykorzystaniem testu chi-kwadrat najmniej różniły się od rezultatów uzyskanych za pomocą testu Andersona-Darlinga, a najbardziej od rezultatów uzyskanych za pomocą testu Kołmogorowa-Smirnowa i Craméra-von Misesa (tab. 1, 2). Porównując testy Kołmogorowa-Smirnowa, Craméra-von Misesa i Andersona-Darlinga, zauważamy, że najbardziej zbliżone rezultaty otrzymano dla testu Kołmogorowa-Smirnowa i Craméra-von Misesa (tab. 1, 2).

Wielkości obciążenia oraz pierwiastka z błędu średniokwadratowego świadczą o dobrym dopasowaniu wybranych rozkładów teoretycznych do rozkładów empirycznych (tab. 3). Wartości znormalizowanego obciążenia mniejsze od 0,0004 i 0,0005 odpowiednio dla rozkładu mieszanego Weibulla i gamma sugerują, że krzywe obu tych rozkładów wyrównują dane empiryczne

Tabela 1.
Wartości statystyk testowych oraz graniczny poziom istotności dla rozkładu mieszanego Weibulla
Test statistics and threshold significance level of the mixture Weibull model

Powierzchnia	Test chi-kwadrat		Test Kołmogorowa-Smirnowa		Test Craméra-von Misesa		Test Andersona-Darlinga	
	χ^2	p	D	p	W^2	p	A^2	p
ŚPN21	38,14	0,4176	0,03	0,9874	0,04	0,9041	0,36	0,8728
ŚPN22	11,72	0,9255	0,04	0,9648	0,04	0,9196	0,38	0,8234
ŚPN23	42,13	0,0027	0,06	0,6073	0,10	0,5679	0,96	0,3572
ŚPN24	33,12	0,4614	0,02	1,0000	0,01	0,9999	0,09	0,9998
ŚPN25	36,23	0,0391	0,08	0,4223	0,14	0,4259	0,73	0,4623
ŚPN26	40,52	0,0594	0,04	0,9982	0,04	0,9215	0,35	0,8887
ŚPN27	42,21	0,0171	0,04	0,9975	0,03	0,9687	0,26	0,9469
ŚPN28	69,78	<0,001	0,07	0,7833	0,07	0,7163	0,60	0,6132
ŚPN29	10,44	0,7915	0,06	0,7120	0,13	0,4650	0,74	0,5170
ŚPN30	22,26	0,3849	0,05	0,9636	0,07	0,7395	0,50	0,6913
ŚPN31	53,95	0,0065	0,05	0,6942	0,08	0,6648	0,64	0,5985
3028	11,77	0,2266	0,07	0,9983	0,03	0,9642	0,27	0,9099
37022	17,62	0,1724	0,06	0,9988	0,02	0,9915	0,21	0,9499
75045	21,88	0,1111	0,14	0,5348	0,11	0,5092	0,83	0,4429
71025	19,63	0,2936	0,05	1,0000	0,02	0,9702	0,18	0,9348
183084	11,73	0,5500	0,04	0,9994	0,01	0,9956	0,14	0,9830
111039	×	×	×	×	×	×	×	×
57049	17,64	0,0614	0,06	0,9992	0,04	0,8941	0,29	0,8843
11025	26,22	0,2905	0,07	0,9978	0,04	0,9340	0,30	0,9132
23057	17,80	0,1654	0,05	0,9920	0,02	0,9820	0,13	0,9898
75028	14,13	0,9595	0,04	0,9999	0,01	1,0000	0,18	0,9811
99043	24,57	0,6513	0,05	1,0000	0,02	0,9874	0,30	0,9175
71013	11,46	0,6500	0,08	0,9451	0,05	0,8626	0,25	0,9249
113079	9,30	0,6767	0,07	0,9886	0,05	0,8554	0,26	0,9061
115045	22,97	0,4625	0,08	0,8967	0,06	0,8008	0,41	0,7926
199016	16,34	0,4292	0,10	0,7909	0,04	0,8948	0,19	0,9677
23046	20,42	0,0255	0,11	0,7080	0,10	0,5866	0,43	0,7410
27013	32,09	0,0097	0,08	0,9668	0,05	0,8794	0,53	0,6605
39065	14,26	0,4305	0,07	0,9769	0,02	0,9701	0,17	0,9664
193028	7,35	0,1956	0,07	0,9864	0,03	0,9507	0,48	0,7280

× – brak zbieżności procedury numerycznej; lack of convergence of numerical procedure

„symetrycznie", czyli zbliżone różnice między danymi empirycznymi a krzywymi teoretycznymi występują „nad i pod" każdą krzywą. Wartości znormalizowanego pierwiastka z błędu średniokwadratowego mniejsze od 0,032 w przypadku obu analizowanych rozkładów pozwalają na stwierdzenie, że różnice między danymi empirycznymi a krzywymi teoretycznymi nie są duże (tab. 3).

Dyskusja

Najczęściej stosowanymi testami zgodności dla jednej próby jest test chi-kwadrat oceniający różnice między empiryczną i teoretyczną liczbą elementów w wyróżnionych przedziałach badanej cechy oraz nieparametryczne testy Kołmogorowa-Smirnowa, Craméra-von Misesa i Andersona-Darlinga, bazujące na analizie różnic między dystrybuantą empiryczną i teoretyczną. Test chi-kwadrat może być stosowany dla cech skokowych i ciągłych, natomiast pozostałe testy w swojej

Tabela 2.
Wartości statystyk testowych oraz graniczny poziom istotności dla rozkładu mieszanego gamma
Test statistics and threshold significance level of the mixture gamma model

Powierzchnia	Test chi-kwadrat		Test Kołmogorowa-Smirnowa		Test Craméra-von Misesa		Test Andersona-Darlinga	
	χ^2	p	D	p	W^2	p	A^2	p
ŚPN21	40,74	0,3094	0,04	0,8575	0,10	0,5997	0,58	0,6587
ŚPN22	12,38	0,9023	0,05	0,9456	0,05	0,8936	0,41	0,7887
ŚPN23	43,61	0,0017	0,07	0,5456	0,11	0,5278	1,08	0,3014
ŚPN24	34,67	0,3884	0,02	1,0000	0,01	0,9996	0,11	0,9991
ŚPN25	33,09	0,0796	0,08	0,5179	0,09	0,6064	0,45	0,7036
ŚPN26	39,16	0,0784	0,05	0,9792	0,05	0,8966	0,33	0,9011
ŚPN27	44,61	0,0093	0,05	0,9589	0,04	0,9040	0,32	0,8923
ŚPN28	70,01	<0,001	0,07	0,6804	0,09	0,6248	0,69	0,5361
ŚPN29	10,91	0,7591	0,07	0,6610	0,12	0,4739	0,76	0,5055
ŚPN30	23,18	0,3344	0,05	0,9112	0,08	0,6522	0,63	0,5712
ŚPN31	51,93	0,0106	0,05	0,7542	0,08	0,6647	0,67	0,5677
3028	12,34	0,1946	0,07	0,9977	0,04	0,9263	0,33	0,8552
37022	17,46	0,1791	0,07	0,9917	0,02	0,9876	0,22	0,9430
75045	22,91	0,0861	0,16	0,3608	0,15	0,3881	0,90	0,4003
71025	19,65	0,2926	0,05	1,0000	0,02	0,9620	0,20	0,9210
183084	11,76	0,5478	0,04	0,9995	0,02	0,9938	0,14	0,9814
111039	20,61	0,8411	0,04	0,9996	0,01	0,9985	0,18	0,9836
57049	15,84	0,1045	0,05	1,0000	0,04	0,9314	0,29	0,8845
11025	27,38	0,2400	0,08	0,9828	0,04	0,9284	0,30	0,9190
23057	18,96	0,1244	0,07	0,9421	0,04	0,9143	0,20	0,9633
75028	14,13	0,9595	0,03	1,0000	0,01	1,0000	0,17	0,9835
99043	24,92	0,6325	0,05	0,9999	0,03	0,9844	0,29	0,9204
71013	11,64	0,6350	0,07	0,9615	0,04	0,8925	0,22	0,9504
113079	8,93	0,7086	0,08	0,9689	0,05	0,8655	0,25	0,9201
115045	22,47	0,4921	0,08	0,9064	0,06	0,7683	0,49	0,7191
199016	15,79	0,4676	0,10	0,7932	0,04	0,9055	0,18	0,9728
23046	20,21	0,0273	0,12	0,7094	0,11	0,5176	0,48	0,6834
27013	32,29	0,0092	0,09	0,9391	0,05	0,8441	0,60	0,6000
39065	14,24	0,4320	0,08	0,9889	0,03	0,9609	0,18	0,9568
193028	8,04	0,1542	0,06	0,9933	0,03	0,9693	0,40	0,7966

klasycznej formie mogą być wykorzystane w przypadku analizy cech ciągłych. Obecnie istnieją wersje testów Kołmogorowa-Smirnowa, Craméra-von Misesa i Andersona-Darlinga opracowane również dla cech skokowych [Arnold, Emerson 2011].

Dana cecha jest cechą ciągłą, jeżeli prawdopodobieństwo wystąpienia w próbie dwóch jednakowych wartości jest równe zeru. W praktyce wielkości badane mierzy się tylko z pewną dokładnością związaną z przyjętym układem jednostek, co prowadzi do grupowania wyników. Jeżeli jednostka przyjętej dla danego zagadnienia skali zwiększa się, to coraz więcej mierzonych elementów ma jednakowe wartości. Opisany mechanizm prowadzi do przekształcenia cechy ciągłej w cechę skokową. Przedstawione zjawisko występuje podczas pomiaru pierśnic. Grubość drzew jest cechą ciągłą, ale przyjęta jednostka, najczęściej 1 lub 0,5 cm, powoduje grupowanie wyników. Następuje przekształcenie cechy ciągłej w cechę skokową. Na badanych powierzchniach maksymalna liczba pierśnic o tej samej wartości kształtowała się od 2 (powierzchnia 75028) do 20 (po-

Tabela 3.

Znormalizowane obciążenie (NB) i znormalizowany pierwiastek z błędu średniokwadratowego (NRMSE) dla badanych modeli

Normalised bias (NB) and normalised root mean square error (NRMSE) of investigated models

Powierzchnia	Rozkład mieszany Weibulla		Rozkład mieszany gamma	
	NB	NRMSE	NB	NRMSE
ŚPN21	0,0001	0,0279	0,0293	0,0001
ŚPN22	0,0005	0,0310	0,0297	0,0010
ŚPN23	0,0005	0,0641	0,0646	0,0010
ŚPN24	0,0005	0,0143	0,0143	0,0006
ŚPN25	0,0017	0,0332	0,0307	0,0010
ŚPN26	0,0002	0,0314	0,0298	0,0004
ŚPN27	0,0007	0,0418	0,0424	0,0011
ŚPN28	0,0002	0,0503	0,0504	0,0005
ŚPN29	0,0009	0,0271	0,0308	0,0009
ŚPN30	0,0004	0,0327	0,0352	0,0005
ŚPN31	0,0004	0,0324	0,0308	0,0005
3028	0,0020	0,0336	0,0370	0,0022
37022	0,0002	0,0301	0,0300	0,0004
75045	0,0002	0,0306	0,0314	0,0004
71025	0,0003	0,0176	0,0174	0,0003
183084	0,0004	0,0172	0,0172	0,0004
111039	×	×	0,0099	0,0001
57049	0,0006	0,0528	0,0475	0,0007
11025	0,0003	0,0212	0,0216	0,0004
23057	0,0008	0,0236	0,0280	0,0007
75028	0,0001	0,0087	0,0087	0,0001
99043	0,0001	0,0106	0,0108	0,0002
71013	0,0004	0,0250	0,0227	0,0005
113079	0,0003	0,0295	0,0276	0,0004
115045	0,0001	0,0131	0,0126	0,0001
199016	0,0004	0,0181	0,0174	0,0005
23046	0,0006	0,0569	0,0550	0,0010
27013	0,0009	0,0327	0,0327	0,0010
39065	0,0005	0,0284	0,0281	0,0004
193028	0,0037	0,0708	0,0737	0,0046

× – brak zbieżności procedury numerycznej; lack of convergence of numerical procedure

wierzchnie ŚPN29 oraz ŚPN30). W tej sytuacji analizowane wartości pierśnic należy traktować jak cechy skokowe.

Test chi-kwadrat, w porównaniu do testów opartych na analizie odpowiednich dystrybuant, posiada dwie bardzo istotne zalety – jest uniwersalny i prosty w użyciu oraz pozwala na estymację parametrów badanych rozkładów teoretycznych z próby [D'Agostino, Stephens 1986]. W praktyce oznacza to, że możemy z danej populacji wylosować próbę, stosując metodę reprezentacyjną, następnie na podstawie wylosowanej próby aproksymować rozkład empiryczny wybranymi rozkładami teoretycznymi, czyli estymować parametry wybranych rozkładów teoretycznych i wreszcie za pomocą testu chi-kwadrat analizować dokładność przeprowadzonych aproksymacji. Otrzymane z próby wyniki można uogólnić dla całej populacji. Testy oparte na analizie odpowiednich dystrybuant cechują się, w porównaniu do testu chi-kwadrat, mniejszym prawdopodobieństwem popełnienia błędów I i II rodzaju, zwłaszcza dla małych prób [D'Agostino, Stephens 1986]. Niestety, nie pozwalają one na estymację parametrów badanych rozkładów teoretycznych z próby i dlatego estymacja parametrów powinna odbywać się z całej populacji. Dla testu Craméra-von Misesa odpowiednie przykłady estymacji parametrów z próby i całej populacji oraz dyskusję różnic między otrzymanymi wynikami dla obu tych wariantów można znaleźć w artykule Lockharta i in. [2007]. W niniejszej pracy test chi-kwadrat, w porównaniu do testów opartych na analizie odpowiednich dystrybuant, wskazywał dodatkowo na duże różnice między rozkładami empirycznymi i teoretycznymi w przypadku 5 powierzchni, co jest spowodowane jego mniejszą mocą [Lockhart i in. 2007]. Analizowane drzewostany stanowią pewne przykładowe populacje i dlatego w tym przypadku możemy zastosować oprócz testu chi-kwadrat również testy Kołmogorowa-Smirnowa, Craméra-von Misesa i Andersona-Darlinga. Test Craméra-von Misesa ma lepsze właściwości niż test Kołmogorowa-Smirnowa, lecz jest stosunkowo nieczuły na odstępstwa rozkładu empirycznego od rozkładu teoretycznego w punktach dalekich od średniej. Jego modyfikacja, test Andersona-Darlinga, jest pozbawiony tej wady.

Oprócz testów zgodności do oceny stopnia dopasowania rozkładu teoretycznego do rozkładu empirycznego stosowane są również inne statystyki. Na ogół wykorzystywane są obciążenie, błąd średniokwadratowy i ewentualnie indeks błędu Reynoldsa. Najwięcej informacji dostarczają testy, ponieważ umożliwiają obliczenie odpowiednich poziomów istotności. Pozostałe statystyki pełnią raczej funkcje pomocnicze, pozwalając np. na ocenę „symetryczności" dopasowania i na przeanalizowanie różnic między danymi empirycznymi a krzywymi teoretycznymi. Ciekawą formę prezentacji wyników stopnia dopasowania zaproponowali Poudel i Cao [2013], wprowadzając nową metodę obliczania rankingu poszczególnych metod oraz prezentując otrzymane rankingi na wykresie antenowym.

Podsumowanie

Wyniki testów zgodności dla jednej próby, wielkości obciążenia oraz pierwiastka z błędu średniokwadratowego świadczą o dobrym dopasowaniu badanych rozkładów mieszanych do rozkładów empirycznych pierśnic.

Podczas wyboru testów zgodności dla jednej próby, pozwalających na ocenę dokładności aproksymacji rozkładów empirycznych pierśnic rozkładami mieszanymi, proponuje się przyjąć następujący sposób postępowania:

- jeżeli badane powierzchnie wchodzą w skład pobranej próby (nie stanowią populacji), to stosujemy test chi-kwadrat;
- jeżeli analizowana powierzchnia stanowi populację i charakter danych wskazuje na cechę ciągłą (dwie, trzy jednakowe wartości pierśnic występują sporadycznie), to stosujemy test Andersona-Darlinga (w wersji dla cechy ciągłej);
- jeżeli analizowana powierzchnia stanowi populację i charakter danych wskazuje na cechę skokową (cztery i więcej jednakowych wartości pierśnic występuje często), to stosujemy test Andersona-Darlinga (w wersji dla cechy skokowej).

Literatura

Anderson T. W., Darling D. A. 1952. Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. Ann. Math. Stat. 23: 193-212.

Arnold T. B., Emerson J. W. 2011. Nonparametric goodness-of-fit tests for discrete null distributions. The R Journal 3/2: 34-39.

Bruchwald A. 2001. Möglichkeiten der Anwendung von Wuchsmodellen in der Praxis der Forsteinrichtung. Beitr. Forstwirtsch. u. Landschaftsökol 3: 118-122.

Bruchwald A., Dmyterko E., Wojtan R. 2011. Model wzrostu dla modrzewia europejskiego (*Larix decidua* Mill.) wykorzystujący cechy taksacyjne drzewostanu. Leś. Pr. Bad. 72: 77-81.

Bruchwald A., Zasada M. 2010. Model wzrostu modrzewia europejskiego (*Larix decidua* Mill.). Sylwan 154 (9): 615--624.

Choulakian V., Lockhart R. A., Stephens M. A. 1994. Cramér-von Mises statistics for discrete distributions. Can. J. Stat. 22: 125-137.

Conover W. J. 1972. A Kolmogorov goodness-of-fit test for discontinuous distributions. J. Am. Stat. Assoc. 67: 591--596.

Cramér H. 1928. On the composition of elementary errors. Skand. Akt. 11: 141-180.

D'Agostino R. B., Stephens M. A. 1986. Goodness-of-fit techniques. Marcel Dekker, New York.

Gądek K. 2000. Lasy. W: Cieśliński S., Kowalkowski A. [red.]. Świętokrzyski Park Narodowy. Przyroda, gospodarka, kultura. Świętokrzyski Park Narodowy, Bodzentyn, Kraków. 349-378.

Gleser L. J. 1985. Exact power of goodness-of-fit tests of Kolmogorov type for discontinuous distributions. J. Am. Stat. Assoc. 80: 954-958.

Gove J. H., Ducey M. J., Leak W. B., Zhang L. 2008. Rotated sigmoid structures in managed uneven-aged northern hardwood stands: a look at the Burr Type III distribution. Forestry 81: 161-176.

Jaworski A., Podlaski R. 2007. Structure and dynamics of selected stands of primeval character in the Pieniny National Park. Dendrobiology 58: 25-42.

Jaworski A., Podlaski R. 2012. Modelling irregular and multimodal tree diameter distributions by finite mixture models: an approach to stand structure characterization. J. For. Res. 17: 79-88.

Korpeľ Š. 1995. Die Urwälder der Westkarpaten. G. Fischer-Verlag, Stuttgart.

Law A. M., Kelton W. D. 2000. Simulation modeling and analysis. Wiley, New York.

Li F., Zhang L., Davis C. J. 2002. Modeling the joint distribution of tree diameters and heights by bivariate generalized beta distribution. For. Sci. 48: 47-58.

Liu C., Zhang L., Davis C. J., Solomon D. S., Gove J. H. 2002. A finite mixture model for characterizing the diameter distribution of mixed-species forest stands. For. Sci. 48: 653-661.

Lockhart R., Spinelli J., Stephens M. 2007. Cramér-von Mises statistics for discrete distributions with unknown parameters. Can. J. Stat. 35: 125-133.

Macdonald P. D. M., Du J. 2004. mixdist: Mixture Distribution Models.

Macdonald P. D. M., Pitcher T. J. 1979. Age-groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. J. Fish. Res. Board. Can. 36: 987-1001.

von Mises R. E. 1928. Wahrscheinlichkeit, Statistik und Wahrheit. Springer, Wien.

Olano J. M., Palmer M. W. 2003. Stand dynamics of an Appalachian old-growth forest during a severe drought episode. For. Ecol. Manage. 174: 139-148.

Podlaski R. 2010a. Diversity of patch structure in Central European forests: are tree diameter distributions in near--natural multilayered *Abies-Fagus* stands heterogeneous? Ecol. Res. 25: 599-608.

Podlaski R. 2010b. Two-component mixture models for diameter distributions in mixed-species, two-age cohort stands. For. Sci. 56: 379-390.

Podlaski R. 2011a. Modelowanie rozkładów pierśnic drzew z wykorzystaniem rozkładów mieszanych. I. Rozkłady mieszane: definicja, charakterystyka, estymacja parametrów. Sylwan 155 (4): 244-252.

Podlaski R. 2011b. Modelowanie rozkładów pierśnic drzew z wykorzystaniem rozkładów mieszanych. II. Aproksymacja rozkładów pierśnic w lasach wielopiętrowych. Sylwan 155 (5): 293-300.

Podlaski R., Roesch F. A. 2013. Aproksymacja rozkładów pierśnic drzew w dwugeneracyjnych drzewostanach za pomocą rozkładów mieszanych. I. Estymacja parametrów. Sylwan 157 (8): 587-596.

Poudel K., Cao Q. V. 2013. Evaluation of methods to predict Weibull parameters for characterizing diameter distributions. For. Sci. 59: 243-252

Reynolds M. R., Burk T., Huang W-H. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. For. Sci. 34: 373-399.

Slakter M. J. 1965. A comparison of the Pearson chi-square and Kolmogorov goodness-of-fit tests with respect to validity. J. Am. Stat. Assoc. 60: 854-858.

Wang M., Rennolls K. 2005. Tree diameter distribution modelling: introducing the logit-logistic distributions. Can. J. For. Res. 35: 1305-1313.

de Wet T., Venter J. 1994. Asymptotic distributions for quadratic forms with applications to tests of fit. Ann. Stat. 2: 380--387.

Zhang L. J., Gove J. H., Liu C., Leak W. B. 2001. A finite mixture of two Weibull distributions for modeling the diameter distributions of rotated-sigmoid, uneven-aged stands. Can. J. For. Res. 31: 1654-1659.

SUMMARY

Approximation of the breast height diameter distribution of two-cohort stands by mixture models. II. Goodness-of-fit tests

The chi-squared test is the best-known goodness-of-fit test, but it is far from ideal and lacks power for small sample sizes. Other popular, more powerful nonparametric tests are: the Kolmogorov--Smirnov test, the Cramér-von Mises test and the Anderson-Darling test. In their classic forms, these goodness-of-fit tests are intended for continuous hypothesized distributions, but they have also been adapted for discrete distributions. The goals of this study are (1) to analyse the accuracy of the approximation of empirical distribution of diameter at breast height (dbh) using two-component mixtures of either the Weibull distribution or the gamma distribution in two--cohort stands, and (2) to discuss the procedure of choosing goodness-of-fit tests. The study plots were located in the Świętokrzyski National Park (central Poland) and in the Southern Appalachian Mountains (eastern USA). The chi-squared test, the extension of the Kolmogorov--Smirnov test to non-continuous null distributions, the generalizations of the Cramér-von Mises test and the Anderson-Darling test to discrete distributions were used.

The results of goodness-of-fit tests, normalised bias and normalised root mean square error, indicate that dbh empirical distributions of two-cohort stands appear to be compatible with the mixture models investigated. The fit of the mixture Weibull model to empirical dbh distributions had a precision similar to that of the mixture gamma model. The following strategy is recommended for evaluating the accuracy of the approximation of empirical dbh distributions: (1) if the plots are a subpopulation we use the chi-squared test, (2) if the plot is a population and the empirical dbh distribution is continuous (two or three equal values of the empirical dbh are rare) we use the classic Anderson-Darling test, (3) if the plot is a population and the empirical dbh distribution is non-continuous (four or more identical dbhs are often) we use the generalization of the Anderson-Darling test to discrete distributions.