SYLWAN nr 3: 40-46, 2003

CHRIS J. CIESZEWSKI, MICHAŁ ZASADA

Wyprowadzanie ogólnych dynamicznych równań bonitacyjnych za pomocą uniwersalnej metody różnic algebraicznych

Derivation of Generic Dynamic Site Equations Using Generalized Algebraic Difference Approach

ABSTRACT

This article describes use of the Generalized Algebraic Difference Approach for derivation of generic dynamic site equations. The generic equations are useful in situations of unspecific expectations with regards to the final model form where one wants to test simultaneously in a single fitting many various models and assumptions about the modeled processes. Equations derived in this work are extremely flexible with various patterns of polymorphisms and asymptotes. Even thought in principle the generic equations have more parameters, those derived with methodology proposed here are very parsimonious.

KEY WORDS

Site index model, site productivity, base-age invariance, dynamic equations

Wprowadzenie

We współczesnym podejściu do modelowania bonitacji stosowane są niemal wyłącznie funkcje trójwymiarowe. Budowane modele opisują zazwyczaj zależność pomiędzy wiekiem (zmienną niezależną), warunkami początkowymi w postaci co najmniej jednej dodatkowej zmiennej reprezentującej intensywność modelowanych procesów [np. García 1983] i zmienną zależną. Zmienną zależną w tych modelach jest najczęściej wysokość, ale może być to również jakakolwiek inna cecha mierzalna, jak grubość, pierśnicowe pole przekroju, miąższość, liczba drzew, itp. Zmienna reprezentująca intensywność procesów jest wyrażana jako wielkość uwikłana w formie obserwacji zmiennej zależnej.

W 1974 roku Bailey i Clutter wprowadzili koncepcję niezależności przewidywań od wyboru wieku bazowego (base-age invariance), w której równanie dynamiczne może dawać wyniki bezpośrednio z jakiejkolwiek pary danych wiek-wysokość bez utraty dokładności przewidywań. W równaniach wzrostu wysokości niezmienne pozostają wielkości modelowane i jako wynik, kształty krzywych wzrostu wysokości. Równania dynamiczne mogą być traktowane jako zależność czterech zmiennych ciągłych, choć opisują one relacje trójwymiarowe, podobnie jak inne równania ze stałym wiekiem bazowym. W równaniach tych jeden z wymiarów jest reprezentowany przez warunek początkowy, który wykorzystuje dwie zmienne uwikłane. Rozwinięcie metody równań algebraicznych – uniwersalna metoda rożnic algebraicz-

CHRIS J. CIESZEWSKI
Warnell School of Forest Resources
University of Georgia,
Athens, GA, 30602, USA
http://www.growthandyield.com/chris/

MICHAŁ ZASADA
Samodzielny Zakład Dendrometrii i Nauki
o Produkcyjności Lasu, Wydział Leśny, SGGW
ul. Rakowiecka 26/30
02-528 Warszawa
e-mail: les_kpl@delta.sggw.waw.pl

nych [np., Cieszewski 1994, Cieszewski i Bailey 2000, Cieszewski i Zasada 2002] pozwala nie tylko na wyprowadzanie równań dymamicznych niezmiennych dla różnych wartości wieku bazowego, lecz również charakteryzujących się jednocześnie polimorfizmem i zmiennymi asymptotami.

Uniwersalna metoda różnic algebraicznych posługuje się teoretyczną zmienną – współczynnikiem intensywności procesu χ. Jest ona wyrazem tych procesów, które są związane z siedliskiem i indywidualną charakterystyką wzrostu i przeżywalności drzew. Służy ona do opisania dowolnych zmiennych rządzących zmianami kształtu krzywych wzrostowych na różnych siedliskach. Zmienna χ może być zarówno zmienną, jak i funkcją dowolnej liczby zmiennych. Ponieważ pomiar współczynnika intensywności procesu jest niemożliwy, jest on zastępowany możliwymi do uzyskania na drodze pomiaru warunkami początkowymi. Dopiero wówczas takie równanie może być wykorzystane w praktyce.

Praktyczne zastosowanie UMRA może być wprowadzone na różnych poziomach zaawansowania i trudności w wyprowadzaniu równań. Równania proste oparte są na prostym przekształceniu równania podstawowego. Równania zaawansowane wymagają znajdowania pierwiastków równania w celu określenia parametrów [Cieszewski i Bailey 2000, Cieszewski i Zasada 2002]. W niniejszej pracy opisane zostały bardziej zaawansowane kategorie równań dynamicznych. Są to równania ogólne, które mają zastosowanie w przypadku braku dostatecznej informacji na temat modelowanych procesów. W pracy zawarto również przykłady połączenia równań ogólnych z innymi metodami statystycznymi, jak na przykład regresją krokową.

Równania ogólne

Równania ogólne mają szczególne zastosowanie w przypadku braku dostatecznej infomacji na temat modelowanych procesów, gdzie oczekuje się, że model musi być dostosowany do bliżej nieokreślonych warunków, a równania ogólne są formułowane przy braku jawnych oczekiwań dotyczących ostatecznej formy modelu. Mogą być one traktowane jako typowy przykład modelowania wykorzystującego równania dynamiczne. Aby szybciej i efektywniej wybrać równanie, modelujący może chcieć objąć szeroki zakres możliwych równań podczas pojedynczej analizy. Doskonały przykład takiej praktyki można znaleźć w pracy Schnute'a (1981).

Równania ogólne są traktowane jako osobna kategoria z powodu potencjalnie dużej liczby parametrów i skomplikowanej formy pochodzącej z wyprowadzania prostego lub złożonego. Równania ogólne powinny być budowane z dużą uwagą, ponieważ bardzo łatwo ulegają one przeparametryzowaniu, niestabilności i trudnościom związanym z estymacją parametrów modelu. Przykład oparty na bazowym równaniu Schumachera [Bailey i Clutter 1974, Cieszewski i Bailey 2001, Cieszewski i Zasada 2002] może charakteryzować się brakiem specyficznych oczekiwań tak w stosunku do asymptot, jak i w stosunku do parametru kształtu, które są jedynymi lub głównymi wyrażeniami opisującymi intensywność wzrostu. Przykład ten może być odpowiedni do wyprowadzenia równania, które może, ale nie musi, mieć asymptoty uzależnione od czynnika siedliskowego i może, ale nie musi, mieć kształty krzywych zmienne na różnych siedliskach. Ponadto efekty te mogą występować w różnych proporcjach. Rozpatrzmy nastepujące proste równanie podstawowe:

$$2\ln Y(t,\chi)=(\chi-\beta_a/t)+(\alpha_p-\chi/t) \qquad [1]$$

Równanie to można przekształcić [1]:

$$\ln Y(t,\chi) = (\alpha + \alpha'\chi) - (\beta + \beta'\chi)/t \qquad [2]$$

gdzie α' i β' to parametry ważące. W przypadku tej ogólnej formy trójwymiarowego równania jawnego można łatwo sprawdzić, że stosunki przekształconych podstawowych równań składowych wzoru [1] (Cieszewski i Bailey 2000, Cieszewski i Zasada 2002), to znaczy:

$$\ln Y(t,\chi) = \chi - \beta_a/t \qquad [3]$$

i

$$\ln Y(t,\chi) = \alpha_p - \chi/t \qquad [4]$$

lub równania Schumachera (1939)

$$\ln Y(t) = \alpha - \beta/t \qquad [5]$$

i równania

$$\ln Y(t,\chi) = \chi - \beta\chi/t \qquad [6]$$

są najlepszymi połączeniami dla każdych danych. Przedstawione równanie umożliwia również sprawdzenie czy równanie [6] powinno rzeczywiście być wprost proporcjonalne, czy też tylko częściowo proporcjonalne do χ. W celu zilustrowania tych trzech możliwych hipotez, równanie [1] może być zapisane jako ważona suma równań [3] i [4]. Prowadzi to do przekształconej wersji równania [2]:

$$\ln Y(t,\chi) = \alpha'(\chi - \beta_a/t) + \beta'(\alpha_p - \chi/t) \qquad [7]$$

gdzie $\alpha_p = \alpha/\text{ß}'$ i $\text{ß}_a = \text{ß}/\alpha'$.

To samo można zapisać jako ważoną sumę równania [5] i [6], która jest również ekwiwalentem równania [2]:

$$\ln Y(t,\chi) = \alpha'\chi(\alpha - \beta/t) + \beta'(\alpha - \beta/t) \qquad [8]$$

i może być zapisana jako liniowe uogólnienie równania [6]:

$$\ln Y(t,\chi) = (\alpha'\chi + \beta')(\alpha - \beta/t) \qquad [9]$$

gdzie: α' i β' są wagami formy anamorficznej i polimorficznej,
$\alpha_p=\alpha/\beta'$, $\beta_a = \beta/\alpha'$ oraz $\alpha' \neq 0$ i $\beta' \neq 0$.

Rozwiązaniem χ w równaniu [2] jest:

$$\chi = \frac{\ln Y - \alpha - \beta/t}{\alpha' - \beta'/t} = \frac{\ln Y_0 - \alpha - \beta/t_0}{\alpha' - \beta'/t_0} \qquad [10]$$

a po zastosowaniu UMRA, to jest użyciu rozwiązania [10] do zastąpienia wartości χ w równaniu

[2], wynikowe ogólne równanie dynamiczne proste oparte na równaniu Schumachera [5] ma następującą postać:

$$\ln Y(t,t_0,Y_0)=\alpha-\frac{\beta}{t}+\frac{\alpha'-\beta'/t}{\alpha'-\beta'/t_0}\left(\ln Y_0-\alpha+\frac{\beta}{t_0}\right) \quad [11]$$

Powstałe równanie ma jednak wyraźnie za dużo parametrów – do tego stopnia, że staje się ono niedefiniowalne. Sytuacją tę można poprawić przez połączenie parametrów α' i β' w jeden parametr. W zależności od tego, który z dwóch parametrów w równaniu [2] jest równy zero, odpowiednie równanie dynamiczne może przybrać jedną z dwóch form:

$$\ln Y(t,t_0,Y_0)=\alpha-\frac{\beta}{t}+\left(\ln Y_0-\alpha+\frac{\beta}{t_0}\right)\frac{1-\gamma/t}{1-\gamma/t_0} \quad [12a]$$

dla α' ≠ 0 lub

$$\ln Y(t,t_0,Y_0)=\alpha-\frac{\beta}{t}+\left(\ln Y_0-\alpha+\frac{\beta}{t_0}\right)\frac{\delta-1/t}{\delta-1/t_0} \quad [12b]$$

dla β' ≠ 0

gdzie: γ=β'/α' i δ=α'/β' i co najmniej jeden z dwóch parametrów jest różny od zera. Jeżeli zarówno α'=0 i β'=0, nie istnieje równanie siedliskowe zdefiniowane przez równanie [2], ale raczej tylko proste dwuwymiarowe równanie z pojedynczą krzywą, która nie ma nic wspólnego z koncepcją bonitacji czy niezmienności dla różnych wieków bazowych. Można w tym przypadku powiedzieć, że albo dane reprezentują tylko jedno siedlisko lub serię różnych siedlisk zawierającą dużo przecięć, lub zniekształceń sprawiających, że jednoznaczne wyróżnienie oddzielnych klas jest niemożliwe.

Testowanie hipotez dotyczących równań [12a] i [12b] może być prowadzone przez proste testy istotności dla różnych parametrów modelu. Niektóre wnioski płynące z tych testów mogą być następujące:

- β' = 0: Równanie jest anamorficzne i ma zmienne asymptoty,
- α' = 0: Równanie jest polimorficzne i ma jedną asymptotę,
- β' ≠ 0 i α' ≠ 0: Równanie jest polimorficzne i ma zmienne asymptoty,
- | α' | << | β' |: Równanie charakteryzuje się stosunkowo silnym polimorfizmem,
- | α' | >> | β' |: Równanie charakteryzuje stosunkowo silne powiązanie ze zmiennymi asymptotami.

Przykładowe ogólne równanie złożone (zaawansowane) może być wyprowadzone na podstawie następującego równania:

$$\ln Y(t,\chi)=\alpha\chi-\frac{\beta/\chi}{t} \quad [13]$$

Uogólnienie powyższego równania złożonego ma następującą postać:

$$\ln Y(t,\chi)=\alpha\chi-\frac{\beta/\chi}{t} \quad [14]$$

Rozwiązanie wymaga znalezienia pierwiastków równania kwadratowego. Jeżeli równanie ma dwa pierwiastki, konieczne jest ostrożne rozważenie, który z nich jest odpowiedni do wbudowania w ostateczne równanie. Wybór ten może zależeć od parametrów modelu, które z kolei zależą od analizowanych danych i zakresu stosowanego wieku. W tym przykładzie lepszy będzie najczęściej pierwiastek rzeczywisty i dodatni, i dlatego lepiej będzie użyć następującej formy:

$$\chi = 0{,}5(\Re_0 - \alpha)/\alpha' \qquad [15]$$

gdzie:

$$\chi = \beta_0/t_0 + \ln Y_0 + \sqrt{(\ln Y_0 - \alpha + \beta/t_0)^2 + 4\gamma/t_0} \qquad [16]$$

Podstawiając pierwiastek równania kwadratowego w miejsce χ do równania [14], uzyskujemy następujące ogólne równanie dynamiczne:

$$\ln Y(t, t_0, Y_0) = \frac{\Re_0 + \alpha}{2} + \frac{2\gamma/t}{\Re_0 - \alpha} - \frac{\beta}{t} \qquad [17]$$

gdzie: $\gamma = \alpha' \beta'$, które jest uogólnieniem zaawansowanego równania dynamicznego przedstawionego przez Cieszewskiego i Baileya [2000] i Cieszewskiego i Zasadę [2002].

Równania dynamiczne a równania regresji krokowej

Opisywana w tym artykule metodyka kładzie nacisk na rolę modelującego w procesie formułowania hipotez, na podstawie których budowane są równania, przed ich końcowym przekształceniem w równania dynamiczne. Formy równań w tym podejściu są determinowane raczej przede wszystkim przez osobę modelującą, niż przez bliżej nieokreślone analizy statystyczne. Jednakże równania wyprowadzane przedstawioną metodą mogą dostarczyć tak dużej elastyczności przewidywań, że to właśnie statystyczne dopasowanie może byc konieczne, żeby określić ostateczną formę równania dynamicznego.

W praktyce dosyć często statystyka, a nie osoba budująca model ustala formę ostatecznego równania. W sytuacjach takich używa się na przykład regresji krokowej (stepwise regression), permutacyjnej i innych rodzajów analizy regresyjnej (liniowych lub nieliniowych). W metodach tych kryteria wyboru modelu zależą od wyników analizy błędów lub innych statystyk. Może się wydawać, że UMRA nie ma zastosowania w takich sytuacjach. Jednakże tak nie jest. Prezentowana metodyka jest przydatna do ulepszenia istniejących równań regresji nawet wówczas, gdy są one formulowane przez regresję krokową czy inne metody. Rozważmy na przykład następujące równanie z czterema parametrami oparte na regresji krokowej:

$$Y(t, S) = \alpha\sqrt{t} + \beta t^2 \ln^{32} t - \gamma \frac{t^{5/2}}{\ln t} + S\delta\sqrt{t} \qquad [18]$$

z rozwiązaniem S dla warunków początkowych t_0, Y_0:

$$S(t, Y) = -\frac{\alpha t_0 \ln t_0 + t_0 \beta t_0^2 \ln^{33} - t_0^{5/2} - \gamma Y_0 \ln t_0}{\delta \ln t_0 \sqrt{t_0}} \qquad [19]$$

prowadzącym do równania dynamicznego o dwóch parametrach:

$$Y(t, t_0, Y_0) = \beta\left(t^2 \ln^{32} t - t_0^{5/2} t_0 \ln^{32} \sqrt{t}\right) + \gamma\left(\frac{t_0^2\sqrt{t}}{\ln t_0} - \frac{t^{5/2}}{\ln t}\right) + Y_0\sqrt{\frac{t}{t_0}} \qquad [20]$$

UMRA została użyta do przekształcenia przedstawionego modelu w równanie dynamiczne, które:

1. Generuje identyczne krzywe, jak te, które są generowane przez równanie [18],
2. Daje wysokość w wieku bazowym równą indeksowi bonitacyjnemu (bonitacji),
3. Może służyć do określania bonitacji i wysokości za pomocą tego samego równania,
4. Umożliwia bezpośrednie wykorzystanie wysokości i wieku zamiast bonitacji z ustalonym wiekiem bazowym, oraz może być łatwo dopasowane i używane z wykorzystaniem dowolnego wieku bazowego.

Wszystkie te ulepszone właściwości są uzyskane przy zmniejszonej o połowę liczbie parametrów i nie stoją w sprzeczności z teorią regresji czy praktyką.

Podsumowanie i wnioski

Niniejsza praca poświęcona jest omówieniu metodyki budowy ogólnych równań siedliskowych opartych na nieobserwowalnych zmiennych siedliskowych. Opisana metodyka jest bardziej elastyczna, niż inne dotychczas stosowane metody używane do tego celu. Wyprowadzone równania mogą być dopasowane do danych z użyciem dowolnej techniki używanej do budowy równań dynamicznych lub o stałym wieku bazowym. Ponadto równania te mogą być używane w sposób spójny z istniejącymi tradycyjnymi równaniami o stałym wieku bazowym. Są one prostsze (mają mniej parametrów) i bardziej elastyczne. Mogą prognozować odpowiednie wysokości, kiedy wiek równy jest wiekowi bazowemu oraz są łatwiejsze do dopasowania przy użyciu niedostatecznych danych lub danych pochodzących z młodych drzewostanów.

Literatura

Bailey R.L., Clutter J.L. 1974. Base age invariant polymorphic site curves. For. Sci. 20: 155-159.

Cieszewski C.J. 1994. Development of a variable density height-growth-model through defining multidimensional height growth spaces. Praca doktorska. University of Alberta, Edmonton, Kanada.

Cieszewski C.J. Bailey R.L., 2000. Generalized Algebraic Difference Approach: A New Methodology for Derivation of Biologically Based Dynamic Site Equations. Forest Science 46: 116-126.

Cieszewski C.J. Zasada M., 2002. Uniwersalna Metoda Różnic Algebraicznych: uogólnione wyprowadzanie dynamicznych równań bonitacyjnych opartych na teoriach biologicznych. Sylwan, 11: 51-62.

García O. 1983. A stochastic differential equation model for the height growth of forest stands. Biometrics, 39: 1059-1072.

Schnute J. 1981. A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 38: 1128-1140.

Schumacher F.X. 1939. A new growth curve and its application to timber yield studies. J. For. 37: 819-820.

SUMMARY

Derivation of Generic Dynamic Site Equations Using Generalized Algebraic Difference Approach

Modern approaches to site-dependent growth and yield modelling involve using almost exclusively three dimensional functions. They describe usually relationships between age (independent variable), modeled phenomena (dependent variable), and at least one additional variable representing intensity of modelled processes in a form of initial conditions. The dependent variable in such models may be height or any other variable of interest, such as diameter, basal area, volume, number of trees, etc. In 1974 Bailey and Clutter introduced the concept of base age invariance, in which height predictions are unaffected by any arbitrary choices off base ages. Their method of model derivation, called algebraic difference approach and its generalization

(Generalized Algebraic Difference Approach) proposed by Cieszewski in 1994 – can be used to derive base age invariant dynamic site equations capable of describing concurrent polymorphism and variable asymptotes. This article describes the use of the Generalized Algebraic Difference Approach for derivation of generic dynamic site equations. The generic equations are useful in situations of unspecific expectations with regards to the final model form where one wants to test simultaneously many various models and assumptions about the modeled phenomena in a single fitting. The generic equations are treated as a separated category due to their relatively high complexity even when derived from simple bases. Equations derived in this work are extremely flexible with various patterns of polymorphisms and asymptotes. Even though in principle the generic equations have more parameters, those derived with the methodology proposed here are very parsimonious. The advocated method emphasizes the importance of the modeller's role in hypotheses formulation during derivation of generic dynamic equations. In general equations' forms in this method are determined by the modeller's understanding of the modelled process rather than by any undetermined statistical analyses. However, the generic equations derived using described method are so flexible, that a statistical fitting may be necessary to determine the final form of the dynamic equation. Finally, the presented methodology can be useful for improvement of existing equations even if they are formulated using stepwise regressions or other traditional *brute force* methods.