

Dubravos urėdijos veiklos proceso rodiklių prognozavimas

Tomas RUZGAS, Rasa CIBULSKYTĖ (KTU)

el. paštas: tomas.ruzgas@ktu.lt, narasyte_rasssiux@yahoo.com

Reziumė. Darbe pateikiami miškų urėdijos veiklos proceso rodiklių prognozavimo modeliai ir programinė įranga, skirta sprendimų priėmimo palengvinimui. Modeliai ir programinė įranga realizuoti naudojant taikomosios statistikos metodus, objektinio programavimo technologiją bei SAS sistemą.

Raktiniai žodžiai: Dubravos urėdija, prognozavimas, laiko eilučių analizė.

1. Įvadas

Priimant atsakingus sprendimus veiklos proceso organizavimo ir valdymo klausimais, kai siekiama racionaliai naudoti miškų išteklius, dažnai naudojama duomenų bazių informacija, apdorota panaudojus statistikos metodus [2]. Daugelyje Vakarų Europos ir JAV miškotvarkos institucijų kaupiama vykdomos veiklos proceso informacija, kuriami matematiniai modeliai ir analizuojami duomenys panaudojant šiuolaikinę statistikos programinę įrangą [2]. Gauti rezultatai naudojami sprendimų priėmimui. Lietuvoje tokių modelių ir specializuotos programinės įrangos dar nėra, nors daugelyje urėdijų informacija apie miškotvarkos procesą kaupiama duomenų bazėse, o analizuojant dažniausiai apsiribojama tik skaitinių charakteristikų skaičiavimu. Tai apsunkina sprendimų priėmimą ir jų pagrindimą sprendžiant ūkinės veiklos organizavimo, miško išteklių atkūrimo ir kitus uždavinius. Todėl aktualus uždavinys – prognozavimo modelių ir metodikų bei atitinkamų programinių priemonių sukūrimas šių problemų sprendimui.

1957 m. įsteigta Dubravos miškų tyrimo stotis, kuri 1998 m. kovo 4 d. įregistruota kaip Valstybės įmonė Dubravos eksperimentinė-mokomoji miškų urėdija. Nors ši įmonė miškų urėdijos veiklą vykdo jau virš 10 metų, tačiau jai pavaldžių girininkijų kirtaviečių duomenis reglamentuota forma pradėjo kaupti tik kiek daugiau nei prieš ketverius metus. Elektroninėje formoje saugoma informacija apie išduotus leidimus kirsti mišką (biržės numeris, naudojimo būdas, kirtimo rūšis, leistas iškirsti plotas ir t.t.), atliktus kirtimus (iškirstas plotas, iškirstas bendras tūris ir t.t.) [6]. Sukurti duomenų analizės ir prognozavimo modelius bei programines priemones, kurių pagalba būtų galima atlikti visapusišką ir kompleksinę Dubravos miškų urėdijos veiklos proceso duomenų bazių analizę, – didelis darbas. Šiame darbe nagrinėjamų uždavinių tikslu buvo:

- įvertinus Dubravos urėdijos duomenų bazių struktūrą ir sistemos SAS galimybes, sukurti ūkinės veiklos proceso rodiklių prognozavimo posistemę;

- sukurti Dubravos urėdijos ūkinės veiklos proceso duomenų prognozavimo modelius, leidžiančius vertinti įvairius rodiklius (leistas iškirsti plotas ir pan.);
- integruoti sukurtus modelius į statistinės analizės posistemę ir sukurti instrumentą, kuris palengvintų sprendimų priėmimą bei jų pagrindimą sprendžiant miškotvarkos proceso organizavimo klausimus.

2. Urėdijos veiklos proceso rodiklių prognozavimo sistema

Kurti informacines sistemas, panaudojant duomenų prognozavimo modelius ir technologijas, yra sudėtingas ir didelių resursų reikalaujantis uždavinys. Pagrindinis reikalavimas tokioms sistemoms: atspindėti realų analizuojamos situacijos vaizdą greitai besikeičiančiose šiuolaikinėse sąlygose ir padėti priimti optimalų sprendimą bei adekvačiai reaguoti į susidariusią padėtį. Įvertinus Dubravos miškų urėdijos duomenų bazių struktūrą, jose saugomą informaciją ir reikalavimus, keliamus šiuolaikinėms informacijos sistemoms, pasiūlyta miškotvarkos proceso duomenų prognozavimo ir analizės posistemės struktūra. Ją sudaro: sąsajos su Dubravos miškų urėdijos duomenų bazėmis posistemė, duomenų paruošimo analizei posistemė, rodiklių statistinio prognozavimo posistemė, sąsajos su vartotoju posistemė.

Sąsajos su Dubravos miškų urėdijos duomenų bazėmis posistemė perkelia reikiamus duomenis iš jai pavaldžių girininkijų duomenų bazių. Duomenų paruošimo statistinei analizei posistemė formuoja duomenų pjūvius ir paruošia duomenų matricą įvairių statistinių prognozavimo metodų taikymui. Statistinio prognozavimo posistemę sudaro miškotvarkos proceso duomenų analizės ir prognozavimo modeliai. Visas šias posistemas apjungia interaktyvios sąsajos su vartotoju posistemė. Ši posistemė sukurta panaudojus objektinio programavimo priemones turinčius SAS sistemos modulius SAS/AF ir SAS/FSP. Statistinės analizės ir prognozavimo modeliai realizuoti panaudojus SAS/STAT (statistinė analizė), SAS/ETS (laiko eilučių analizė) ir SAS/GRAPH (grafinis duomenų vaizdavimas) modulius.

3. Statistiniai prognozavimo modeliai

Ūkinė veikla, susijusi su miško darbais, labai priklauso nuo sezoninės klimato kaitos. Žinoma, jog miško kirtimo, jo atkūrimo ar selekcijos ir sėklininkystės darbus iš esmės įtakoja metų laikai, todėl atlikus Dubravos urėdijos duomenų bazėse saugomų duomenų analizę pasiūlyti šie statistiniai laiko eilučių modeliai:

1. Sezoniniai $ARIMA(p, d, q)(P, D, Q)_s$ modeliai, kuriuos galima užrašyti:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^d(Y_t - \mu) = \theta(B)\Theta(B^s)\varepsilon_t,$$

čia Y_t – stebėta priklausomo kintamojo reikšmė laiko momentu t , B – poslinkio atgal per laiko vienetą operatorius, t.y. $BY_t = Y_{t-1}$, s – proceso periodiškumas, μ – laiko eilutės vidurkis, ε_t – atsitiktiniai nuokrypiai, o

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

$$\begin{aligned}\Phi(B^s) &= 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}, \\ \Theta(B^s) &= 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}\end{aligned}$$

yra atitinkamai p , q , P ir Q laipsnio daugianariai [1].

2. Sezoniniai eksponentinio glodinimo modeliai:

$$Y_t = \mu_t + s_p(t) + \varepsilon_t,$$

čia μ_t apibūdina vidurkio narį. Funkcija $s_p(t)$ susieja laiką t su vienu iš p sezoninių faktorių. Įprastinėje regresijos struktūroje $s_p(v) = s_p(v + p)$, kai tiesinių dedamųjų visi nariai yra pastovūs. Sezoninio glodinimo modelio lygčių sudarymas iliustruoja lygčių taikymą su dviem glodinimo svoriais:

$$\widehat{Y}_{t+k} = L_t + S_{t-p+k},$$

čia $L_t = \omega(Y_t - S_{t-p}) + (1 - \omega)L_{t-1}$ yra lygmens komponentė su svoriu ω ir $S_t = \delta(Y_t - L_t) + (1 - \delta)S_{t-p}$ yra sezoninė komponentė su svoriu δ [3]. Pradinės reikšmės gaunamos parinktu modeliu prognozuojant eilutę atvirkštine tvarka.

3. Winters adityviniai modeliai, kurie susieja tiesinio trendo modelį su sezoninėmis komponentėmis:

$$Y_t = \mu_t + \beta_t t + s_p(t) + \varepsilon_t,$$

čia β_t apibūdina nuolydį. Priklausomai nuo glodinimo modelio, būseną momentu t susideda iš: L_t glodinto lygmens vertinančio μ_t ; T_t glodinto trendo vertinančio β_t ; S_{t-j} , $j = 1, \dots, p$, sezoninių faktorių vertinančių $s_p(t)$ [7]. Panašiai kaip ir eksponentinio glodinimo atveju naudojamos rekurentinės lygtys:

$$\begin{aligned}L_t &= \omega(Y_t - S_{t-p}) + (1 - \omega)(L_{t-1} + T_{t-1}), \\ T_t &= \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}, \\ S_t &= \delta(Y_t - L_t) + (1 - \delta)S_{t-p},\end{aligned}$$

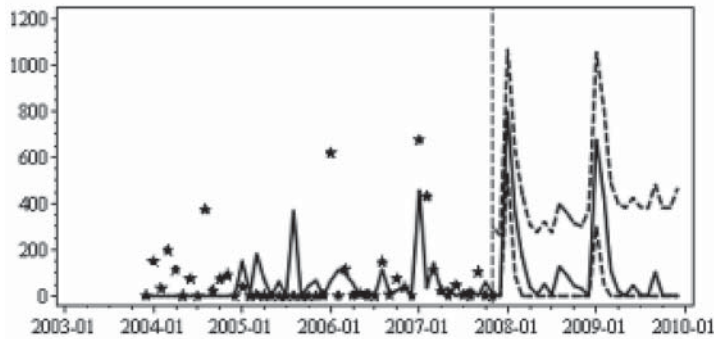
čia γ yra trendo glodinimo svoris. Prognozavimo per k žingsnių lygtis yra:

$$\widehat{Y}_{t+k} = L_t + kT_t + S_{t-p+k}.$$

Pateiksime du prognozavimo modelių taikymo pavyzdžius.

Panaudojus ARIMA(p, d, q)(P, D, Q) $_s$, atliktas leisto iškirsti ploto (ha) prognozavimas. Buvo analizuoti 2003 metų gruodžio–2007 metų lapkričio mėnesių Šilėnų girininkijos kirtaviečių suvestiniai žiniaraščiai. Laiko eilutės stacionarumui ir baltam triukšmui patikrinti naudoti Dickey–Fuller [4] bei Ljung–Box χ^2 [5] kriterijai parodė, jog modelis taikomas korektiškai (hipotezių tikrinimui reikšmingumo lygmuo parinktas $\alpha = 0,05$). Pasinaudojus autokoreliacijų bei dalinių autokoreliacijų funkcijomis su reikšmingais nariais [1,2], sudarytas modelis ARIMA(1, 0, 1)(1, 1, 1) $_{12}$. Įvertinus parametrus, modelį galima užrašyti:

$$(1 - 0,5442B)(1 + 0,9997B^{12})(1 - B^{12})Y_t = (1 - 0,3528B)(1 + 0,9831B^{12})\varepsilon_t.$$



1 pav. Leisto iškirsti ploto (ha) prognozuojama laiko eilutė.

Modelio vidutinės kvadratinės paklaidos šaknis $RMSE=144$. Leisto iškirsti ploto prognozė 2008–2009 metams pateikta 1 pav. Grafike ištisine linija pavaizduota prognozuojama laiko eilutė, o punktyrine – jos pasikliautinasis intervalas su 0,95 pasikliovimo lygmeniu, žvaigždutėmis pažymėtos stebėtos reikšmės.

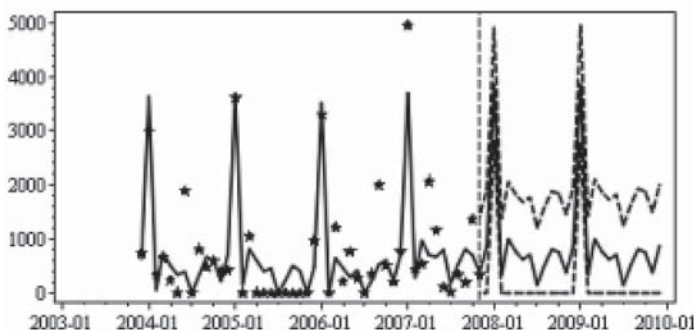
Analizuojant 2003 metų gruodžio–2007 metų lapkričio mėnesių Šilėnų girininkijos kirtavietėse leistą iškirsti tūrį (m^3), patikrinta hipotezė apie baltą triukšmą. Ljung - Box χ^2 kriterijus parodė, jog tikslinga taikyti glodinimo metodus. Panaudojus Winters adityvinį metodą, sudarytas prognozavimo modelis. Modelio glodinimo svoriai yra $\omega = 0,045$, $\delta = 0,001$ ir $\gamma = 0,001$, o rekurentinių lygčių sistemos komponentų pradiniai įverčiai pateikti 1 lentelėje.

Modelio $RMSE=511$, o $R^2 = 0,776$. Leisto iškirsti tūrio prognozė 2008–2009 metams pateikta 2 pav. (žymėjimai naudoti tie patys kaip ir 1 pav.).

Sukurti modeliai ir programinės priemonės analizuoja duomenų bazes įvairiais pjūviais ir gali būti naudojami atliekant ūkinės veiklos proceso statistinį prognozavimą, tiriant atskirų girininkijų ar eiguvų, poūkių ir kirtimo rūšių įtaką miškotvarkos sėkmei, planuojant miškų atkūrimo darbus ir sprendžiant kitus uždavinius.

1 lentelė. Winters modelio parametrų įverčiai

Parametras	Įvertis	Parametras	Įvertis
Glogintas lygmuo L_0	980	Glod. sezoninis faktorius 6	-191
Glodintas trentas T_0	9.3	Glod. sezoninis faktorius 7	-767
Glod. sezoninis faktorius 1	2999	Glod. sezoninis faktorius 8	-406
Glod. sezoninis faktorius 2	-530	Glod. sezoninis faktorius 9	-122
Glod. sezoninis faktorius 3	130	Glod. sezoninis faktorius 10	-181
Glod. sezoninis faktorius 4	-112	Glod. sezoninis faktorius 11	-575
Glod. sezoninis faktorius 5	-269	Glod. sezoninis faktorius 12	28



2 pav. Leisto iškirsti tūrio (m^3) prognozuojama laiko eilutė.

4. Išvados

1. Atlikta Dubravos urėdijos ūkinės veiklos proceso duomenų bazių analizė, pasiūlyta jų statistinės analizės sistemos struktūra, kuri realizuota programiškai panaudojus SAS objektinio programavimo priemones.
2. Panaudojus taikomosios statistikos metodus, sukurti Dubravos urėdijos ūkinės veiklos proceso rodiklių prognozavimo modeliai, kurie leidžia tirti jų ryšius ir priklausomybes laike, o rezultatus panaudoti sprendimų priėmimui.
3. Dalis sukurtų prognozavimo modelių realizuota programiškai. Sukurtos statistinės analizės sistemos architektūra yra atvira, todėl keičiant modelius ar įvedant naujus, ją galima pritaikyti kitų uždavinių sprendimui.

Literatūra

1. G.E.P. Box, G.M. Jenkins, G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd Edition, Prentice Hall, Englewood Cliffs, NJ (1994).
2. J. Brocklebank, D. Dickey, *SAS for Forecasting Time Series*, 2nd edition, SAS Press and John Wiley Sons Inc. (2003).
3. R.G. Brown, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, New York (1962).
4. D.A. Dickey, D.P. Hasza, W.A. Fuller, Testing for unit roots in seasonal time series, *Journal of the American Statistical Association*, **79**(386), 355–367 (1984).
5. G.M. Ljung, G.E.P. Box, On a measure of lack of fit in time series models, *Biometrika*, **65**, 297–303 (1978).
6. *Valstybės įmonės Dubravos eksperimentinės-mokomosios miškų urėdijos įstatai* (2004).
7. P.R. Winters, Forecasting sales by exponentially weighted moving averages, *Management Science*, **6**, 324–342 (1960).

SUMMARY

T. Ruzgas, R. Cibulskytė. Forecasting of activity process indexes for Dubrava forestry

The forecasting models of activity process indexes and statistical software SAS for support the decision-making is presented in this paper. The models and software were developed using applied statistics methods, object programming technology.

Keywords: Dubrava forestry, forecasting, time series analysis.