



KAUNO TECHNOLOGIJOS UNIVERSITETAS
FUNDAMENTALIŲJŲ MOKSLŲ FAKULTETAS
MATEMATINĖS SISTEMOTYROS KATEDRA

Julija Smirnova

VANDENS LYGIO NEMUNO ŽEMUPIO
DELTOJE POTVYNIŲ METU VERTINIMAS

Magistro darbas

Vadovė
doc. dr. V. Karpickaitė

KAUNAS, 2004



KAUNO TECHNOLOGIJOS UNIVERSITETAS
FUNDAMENTALIŲJŲ MOKSLŲ FAKULTETAS
MATEMATINĖS SISTEMOTYROS KATEDRA

TVIRTINU
Katedros vedėjas

prof. habil.dr. V.Pekarskas
2004 06 11

VANDENS LYGIO NEMUNO ŽEMUPIO
DELTOJE POTVYNIŲ METU VERTINIMAS

Taikomosios matematikos magistro baigiamasis darbas

Kalbos konsultantas

dr. I. Džežulskienė

2004 05 31

Recenzentas

doc.dr.J. Venclovienė

2004 06 02

Vadovė

doc. dr. V. Karpickaitė

2004 06 03

Atliko

FMMM - 2gr. stud.

J. Smirnova

2004 05 27

KAUNAS, 2004

KVALIFIKACINĖ KOMISIJA

Pirmininkas: Leonas Saulis, profesorius (VGTU)

Sekretorius: Eimutis Valakevičius, docentas (KTU)

Nariai: Algimantas Jonas Aksomaitis, profesorius (KTU)

Vytautas Janilionis, docentas (KTU)

Vidmantas Povilas Pekarskas, profesorius (KTU)

Rimantas Rudzkis, profesorius (MII)

Zenonas Navickas, profesorius (KTU)

Smirnova J. Water level at Nemunas river delta estimation during floods: Master' s work in applied Mathematics / supervisor doc. Dr. V. Karpickaitė; Department of Mathematics Research In System, Faculty of Fundamental Sciences, Kaunas University of Technology. – Kaunas, 2004. – 73 p.

SUMMARY

Delta of Nemunas River covers an area from the outlet of the tributary Jūra to the Curonian Lagoon in Klaipėda District. In fact, every year large area is inundated during the spring floods.

The main purpose of this paper was to apply statistical methods for extreme events and to estimate probabilities of exceedance of the highest water level recorded and return period of a very large inundated area. A program, using object-programming pack C++ Builder version 5.0, for an easy estimating of these probabilities was created.

Since Nemunas River annual maximum water levels at three gauging – stations – Smalininkai, Tilžė and Rusnė, follow each other time – wise, they were examined in time series context. For each gauging – station water level with a SAS/ETS and SAS Times Series Forecasting System help were created several time series models and selected the one, that fit best.

TURINYS

Įžanga	9
1. Teorinė dalis	11
1.1 Deltos užliejimai	11
1.2 Statistiniai ekstremaliųjų įvykių analizės metodai	12
1.2.1 Momentiniai įrašai (rekordai)	12
1.2.2 Gumbelo perviršių metodas	13
1.2.3 Pasikartojimo periodas	13
1.3 Laiko eilutės	15
1.3.1 Pagrindinės sąvokos	15
1.3.2 Statistinė atsitiktinių sekų analizė ir jų modeliai	17
1.3.3 Stacionarių sekų sąvokos	18
1.3.4 Atsitiktinių sekų modeliai, jų operatorinė forma	19
1.3.5 Stacionarūs tiesiniai modeliai	20
1.3.6 Nestacionarūs tiesiniai modeliai	21
1.3.7 Bendra tiesinio prognozavimo teorija ir statistinių modelių panaudojimas prognozei	22
1.3.7.1 Stacionarus tiesinis procesas	22
1.3.7.2 Nestacionarus tiesinis procesas	23
1.3.8 Modelio eilės nustatymo būdai	24
1.3.8.1 Mažiausios kanoninės koreliacijos metodas (SCAN)	24
1.3.8.2 Išplėtosios imties autokoreliacinės funkcijos metodas (ESACF)	25
2. Tiriamoji dalis	27
2.1 Kintamųjų apibūdinimas ir bendroji analizė	27
2.2 Ekstremaliųjų įvykių statistinių metodų vandens lygiams taikymas	28
2.2.1 Rekordiniai momentai	28
2.2.2 Pasikartojimo periodas	29
2.2.3 Perviršio tikimybės	31
2.3 Kintamojo <i>rusne</i> laiko eilučių modelio parinkimas	32
3. Programinė realizacija ir instrukcija vartotojui	40
Išvados	42
Literatūros sąrašas	43
1 priedas. Rekordiniai momentai	44
2 priedas. Pasikartojimo tikimybės	45
3 priedas. Perviršių tikimybės	46

	6
4 priedas. Laiko eilučių modelio kintamajam <i>smalininkai</i> parinkimas.....	47
5 priedas. Programų tekstai	53
5.1 Tikimybių skaičiavimas	53
5.2 Laiko eilučių modelių sudarymas	63
6 priedas. Straipsniai	65
6.1 Statistinių metodų taikymas ekstremaliesiems įvykiams – potvyniams Nemuno deltoje – tirti..	65
6.2 Vandens lygio Nemuno žemupio deltoje potvynių metu įvertinimas	72

LENTELIŲ SĄRAŠAS

1.1 lentelė. Aukščiausieji pavasario potvynių vandens lygiai Nemuno žemupyje	12
2.1 lentelė. Analizuojamų kintamųjų imčių statistikos	29
2.2 lentelė. Didžiausių Šilutės rajono potvynių suvestinė	30
2.3 lentelė. Tikimybė bent vieną kartą pasirodyti įvykiui “ypatingai didelis potvynis” <i>k</i> metų laikotarpyje	31
2.4 lentelė. 5 metų laikotarpio perviršių tikimybės	32
2.5 lentelė. Baltojo triukšmo hipotezės tikrinimas kintamojo <i>rusne</i> procesui	33
2.6 lentelė. Baltojo triukšmo hipotezės tikrinimas ARIMA(2,1,0) modelio liekanoms	36
2.7 lentelė. Baltojo triukšmo hipotezės tikrinimas MA(4) modelio liekanoms	36
2.8 lentelė. Baltojo triukšmo hipotezės tikrinimas AR(2) modelio liekanoms	36
2.9 lentelė. Baltojo triukšmo hipotezės tikrinimas ARIMA(4,1,1) modelio liekanoms	38
2.10 lentelė. Kintamojo <i>rusne</i> sudarytų alternatyvių modelių paklaidų palyginimas	38
2.11 lentelė. Modelio AR(2) sudarytų prognozių patikrinimas	39
PRIEDAI	
2.1 lentelė. Tikimybė bent vieną kartą <i>k</i> metų laikotarpyje potvynio metu užtvindyti 45 000 ha teritoriją	46
3.1 lentelė. 10 metų laikotarpio perviršių tikimybės	47
4.1 lentelė. Baltojo triukšmo hipotezės tikrinimas AR(1) modelio liekanoms	48
4.2 lentelė. Baltojo triukšmo hipotezės tikrinimas MA(4) modelio liekanoms	50
4.3 lentelė. Baltojo triukšmo hipotezės tikrinimas ARIMA(4,1,0) modelio liekanoms	51
4.4 lentelė. Kintamojo <i>smalininkai</i> sudarytų alternatyvių modelių paklaidų palyginimas	52
4.5 lentelė. Modelio AR(1) sudarytų prognozių patikrinimas	53

PAVEIKSLŲ SĄRAŠAS

2.1 pav. Kintamųjų stebėjimų imtys	28
2.2 pav. Kintamojo <i>rusne</i> rekordiniai momentai	30
2.3 pav. Apskaičiuotos įvykio “ypatingai didelis potvynis” pasirodymo tikimybės	32
2.4 pav. 5 metų laikotarpio perviršių tikimybės	33
2.5 pav. Dickey – Fuller’io testas $(p+d,q)=(1,0)$	34
2.6 pav. AR(1) modelis kintamajam <i>rusne</i>	35
2.7 pav. AR(2) modelis kintamajam <i>rusne</i>	37
2.8 pav. Dickey – Fuller’io testas $(p+d,q)=(5,1)$ procesui	37
2.9 pav. AR(2) modelis kintamajam <i>rusne</i>	39
2.10 pav. Prognozavimo paklaidos, apskaičiuotos kintamajam <i>rusne</i> , pritaikius AR(2) modelį	40
3.1 pav. Programos langas	42
 PRIEDAI	
1.1 pav. Kintamojo <i>tilze</i> rekordiniai momentai	45
1.2 pav. Kintamojo <i>smalininkai</i> rekordiniai momentai	45
2.1 pav. Tikimybė, kad bent vieną kartą per stebimą laikotarpį potvynio metu bus užtvindyta 45 000 ha teritorija	46
3.1 pav. 10 metų laikotarpio perviršių tikimybės	47
4.1 pav. Dickey – Fuller’io testas $(p+d, q)=(1,0)$ procesui	48
4.2 pav. AR(1) modelis kintamajam <i>smalininkai</i>	49
4.3 pav. Dickey – Fuller’io testas $(p+d,q)=(5+0,0)$ procesui	50
4.4 pav. ARIMA(4,1,0) modelis kintamajam <i>smalininkai</i>	51
4.5 pav. AR(1) modelis kintamajam <i>smalininkai</i>	52
4.6 pav. Prognozavimo paklaidos, apskaičiuotos kintamajam <i>smalininkai</i> pritaikius AR(1) modelį	53

IŽANGA

Gamtinės nelaimės, atnešančios didžiulius nuostolius gyventojams, vadinamos ekstremaliais įvykiais. Daugiausia nuostolių visame pasaulyje atnešančios gamtinės stichijos – žemės drebėjimai – Lietuvai negresia. Ekstremalių įvykių pavyzdžiu Lietuvoje yra potvyniai Nemuno deltoje ir prie jos prigludusioje žemupio dalyje. 605 km² plote plyti vos daugiau nei metru aukščiau jūros lygio pakilusi lyguma, kurią visą arba jos dalį kasmet, o kartais ir dažniau, užlieja vanduo, atnešantis didelių nuostolių žmonių turtui, o kartais ir sukelia pavojų žmonių gyvybei.

Labai svarbiu faktoriumi laikomas potvynio metu kylančio vandens lygis. Kuo jis didesnis, tuo didesni plotai gali būti apsemti, tuo daugiau žalos gali būti padaryta. Iš kitos pusės, kylantis vanduo atneša dumblą, kuris tręšia pievas ir laukus, tokiu būdu atnešdamas naudą žemdirbystei. Tačiau, atnešama žala yra didesnė. Siekiant sumažinti nuostolius jau nuo seno vandens lygiai potvynio metu buvo stebimi ir fiksuojami, tačiau apie 1990 metus šis darbas nutrūko. Prieš keletą metų Kėdainių vandens ūkio institutas šiuos stebėjimus atnaujino.

Hidrologinių reiškinių stebėjimų rezultatai pasižymi daugiau arba mažiau nereguliais svyravimais. Jeigu stebimi dydžiai išsidėsto vienas paskui kitą laiko atžvilgiu, tai gaunama seka, kuri nagrinėjama kaip laiko eilutė. Pokyčių dėsningumas tokioje eilutėje turi tendenciją mažėti didėjant laiko tarpui tarp stebėjimų. Pavyzdžiui, vidutinės paros nuotėkio reikšmės nedaug skiriasi viena nuo kitos, o vidutinės metinės nuotėkio reikšmės žymiai skiriasi. Kadangi šioms sekoms būdingas nereguliarumo bruožas, tai analizė geriausiai atliekama taikant statistinius metodus. Ypač tuo atveju, kai nedėsningi svyravimai yra vyraujantys.

Vienas iš svarbiausių potvynio vertinimo uždavinių – nustatyti potvynio dydį, kad būtų galima sumažinti riziką įvairių hidraulinių konstrukcijų (užtvankos, dambos, tiltai ir t.t.) ir žmonių gyvybei.

Yra du būdai nustatyti didžiausius potvynius: determinuotas ir tikimybinis. Tikimybinis būdas yra paremtas paskutiniųjų potvynių dažnio analize ir reikalauja mažiau duomenų nei determinuotasis. Nagrinėjant potvynius tokia statistinė analizė atliekama nebuvo. Šis darbas pradėtas bendradarbiaujant su vandens ūkio institutu, esančiu Kėdainiuose.

Darbe nagrinėjami vandens lygio stebėjimai, užfiksuoti trijuose Nemuno žemupio matavimo postuose – Smalininkai, Tilžė ir Rusnė. Analizei naudojamos tik metinės didžiausios vandens lygių reikšmės.

Pirmoje darbo dalyje taikomi statistiniai metodai, kurie naudojami užsienyje ekstremaliems įvykiams vertinti, apskaičiuojamos perviršių tikimybės, įvertinamas didžiausio potvynio pasikartojimo periodas. Šia tema Lietuvos žemės ūkio universiteto Vandens ūkio instituto žurnalo “Vandens ūkio inžinerija” 22(44) numeryje buvo išspausdintas straipsnis “Statistinių metodų taikymas ekstremaliems įvykiams – potvyniams Nemuno deltoje - tirti”.

Antroje dalyje vandens lygiai nagrinėjami kaip laiko eilutės. Kiekvienam stebėjimo postui buvo sudaryti keli laiko eilučių modeliai ir parinktas geriausiai atitinkantis tikrovę. Šia tema buvo skaitytas pranešimas “Vandens lygio Nemuno žemupio deltoje potvynių metu vertinimas” V studentų konferencijoje, vykusioje Kauno Technologijos Universitete 2004 metų gegužės mėnesį. Analizė atlikta naudojant SAS statistinio paketo ekonometrikos ir laiko eilučių analizės posistemę (SAS/ETS) ir laiko eilučių prognozavimo posistemas (SAS Time Series Forecasting System ir Time Series Viewer).

1. TEORINĖ DALIS

1.1 DELTOS UŽLIEJIMAI

Nemuno žemupyje užliejama zona apima teritoriją nuo Rambyno aukštumos iki Kuršių marių. Kadangi didžioji užliejamos teritorijos dalis yra deltoje, ją galima vadinti Nemuno deltos užliejama teritorija. Yra nustatyta, kad bendras užliejamos teritorijos plotas apima 605 km², iš jų 402 km² tenka dešiniojo kranto deltai, priklausančiai Lietuvai.

Pavasario potvynis yra vandeningoji upės hidrologinio režimo fazė, pasižyminti staigiu ir aukštu vandens lygio pakilimu, dideliais debetais ir ilgesniu bei lėtesniu vandens lygio ir debetų žemėjimu.

Potvyniai Nemuno deltoje stebimi nuo 1901 m. Kaip rodo Smalininkų bei Rusnės vandens matavimo postuose atlikti stebėjimai, aukščiausieji pavasario potvyniai ir didžiausieji deltos užliejimai buvo 1886, 1888, 1917, 1951 ir 1958 metais, jie parodyti 1.1 lentelėje.

1.1 lentelė

Aukščiausieji pavasario potvynių vandens lygiai Nemuno žemupyje

Metai	Vandens lygis matavimo postuose, m	
	Smalininkai	Rusnė
1986	15,25	2,84
1988	15,07	3,50
1917	15,25	2,84
1951	15,56	2,68
1958	16,15	2,85

Deltos užliejimo procesą sąlygoja pastovūs (paviršiaus mikroreljefas, senvagių jungimasis su Nemunu bei jo šakomis) ir kintantys faktoriai (didžiausias vandens lygio aukštis ir trukmė).

Išskiriamos trys užliejimo fazės. Pirmoji – vandens išsiliejimas iš vagos. Pradžioje vanduo iš krantų ima lieti žemiausiose vietose, per ištakų bei senvagių žiotis, vėliau jau ir per aukštesnes kranto vietas. Apsemti Nemuno deltos plotai kuri laiką dar būna izoliuoti, vanduo juose teka ratu arba beveik stovi. Kylant vandens lygiui, po kurio laiko, atskiri užlieti plotai vienas su kitu susijungia ir atsiranda ištisinės tranzitinės srovės. Ši fazė prasideda vidutiniškai kovo mėnesio viduryje ir trunka 2 – 3 dienas, per kurias vandens lygis užlietose pievose pakyla vidutiniškai 1 – 1,5 m.

Antroji užliejimo fazė apima laikotarpį kol egzistuoja šis tranzitinis srautas. Ji trunka 3 – 4 savaites.

Trečioji fazė – tranzitinio srauto nutrūkimas vandens lygiui slūgstant. Pradėjus potvyniui slūgti, vėl atsiranda atskiri, vienas nuo kito izoliuoti, apsemti ruožai, kurių plotas palaipsniui mažėja. Antroji fazė pereina į trečiąją gana neryškiai, todėl tiksliai nustatyti jų trukmę sunku.

1.2 STATISTINIAI EKSTREMALIŲJŲ ĮVYKIŲ ANALIZĖS METODAI

Čia pristatomi keli analitiniai ekstremaliųjų įvykių tyrimo metodai, kurie suteikia naudingą preliminarią informaciją apie nagrinėjamus duomenis.

1.2.1 MOMENTINIAI ĮRAŠAI (REKORDAI)

Tarkime, turime nepriklausomus vienodai pasiskirsčiusius atsitiktinius dydžius X_i su pasiskirstymo funkcija F .

Imkime atsitiktinių dydžių, įgyjančių sveikas teigiamas reikšmes, seką N_1, N_2, \dots, N_n .

n -tuju rekordiniu momentu vadinamas atsitiktinis dydis N_n , apibrėžtas tokiu būdu:

$$N_n = \min \{ j : j > N_{n-1}, X_j > X_{N_{n-1}} \}, n \geq 2 \quad (1.2.1)$$

Įrašas (rekordas) X_n fiksuojamas, jei $X_n > N_{n-1} = \max(X_1, \dots, X_{n-1})$.

X_1 apibrėžiamas kaip pirmasis įrašas (rekordas). Įrašo (rekordo) momento laikas L_n yra atsitiktinis laikas, kai pasikeičia (N_n) reikšmė.

Įrašų (rekordų) skaičiavimo procesas:

$$N_1 = 1, \quad N_n = 1 + \sum_{k=2}^n I_{\{X_k > M_{k-1}\}}, \quad n \geq 2. \quad (1.2.2)$$

Tarkime (X_i) - nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai su tolydžia pasiskirstymo funkcija F ir įrašų (rekordų) skaičiavimo procesu N_n , kuris buvo apibrėžtas anksčiau. Tada

$$EN_n = \sum_{k=1}^n \frac{1}{k} \quad \text{ir} \quad \text{var}(N_n) = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k^2} \right).$$

Iš N_n apibrėžimo:

$$EN_n = 1 + \sum_{k=2}^n P(X_k > M_{k-1}) = 1 + \sum_{k=2}^n \int_{-\infty}^{+\infty} P(X_k > u) dP(M_{k-1} \leq u).$$

Pasinaudoję $P(M_{k-1} \leq u) = F^{k-1}(u)$, gauname EN_n . Analogiškai ir $\text{var}(N_n)$.

EN_n ir $\text{var}(N_n)$ abu yra $\ln n$ eilės, kai $n \rightarrow \infty$.

Tiksliau, $EN_n - \ln n \rightarrow \gamma$, kur $\gamma = 0.5772\dots$ - Eulerio konstanta.

Išvada: nepriklausomų vienodai pasiskirsčiusių dydžių įrašų (rekordų) skaičius auga labai lėtai.

1.2.2 GUMBELO PERVIRŠIŲ METODAS

Gumbelo perviršių metodas atsako į klausimą: kokios reikšmės iš būsimų stebėjimų viršys praeities stebėjimus.

Tegul $X_{n,n} < \dots < X_{1,n}$ - sutvarkyta imties X_1, \dots, X_n statistika, turinti tolydžią pasiskirstymo funkciją F . Imame k -tos viršutinės eilės statistiką $X_{k,n}$ kaip atsitiktinio slenksčio reikšmę ir pažymime $S_r^n(k)$, čia $r \geq 1$ - $X_{k,n}$ perviršių skaičius tarp sekančių r stebėjimų X_{n+1}, \dots, X_{n+r} , t.y.

$$S_r^n(k) = \sum_{i=1}^r I_{\{X_{n+i} > X_{k,n}\}}.$$

Paprastumo dėlei $S_r^n(k)$ žymėsime S .

Atsitiktinis dydis S turi hipergeometrinį skirstinį, t.y.

$$P(S = j) = \frac{\binom{r+n-k-j}{n-k} \binom{j+k-1}{k-1}}{\binom{r+n}{n}}, j = 0, 1, \dots, r. \quad (1.2.3)$$

Kadangi

$$P(S = j) = \int_0^{\infty} P(S = j | X_{k,n} = u) dF_{k,n}(u),$$

čia $F_{k,n}$ - pasiskirstymo funkcija $X_{k,n}$. Pasinaudoję prielaida, kad (X_1, \dots, X_n) ir $(X_{n+1}, \dots, X_{n+r})$ yra nepriklausomi ir $\sum_{i=1}^r I_{\{X_i > u\}}$ turi binominį skirstinį su parametrais r ir $\bar{F}(u)$, iš formulės

$$dF_{k,n}(u) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(u) \bar{F}^{k-1}(u) dF(u)$$

gauname (1.2.3).

Pastaba. Iš S apibrėžimo ir iš aukščiau pateikto argumento išplaukia, kad $ES = \frac{rk}{n+1}$.

1.2.3 PASIKARTOJIMO PERIODAS

Šis metodas atsako į klausimą: kokia tikimybė, kad per tam tikrą laikotarpį pasikartos konkretus ekstremalus įvykis?

Tegul $\{X_i\}$ yra nepriklausomų vienodai pasiskirsčiusių atsitiktinių dydžių seka su tolydžia pasiskirstymo funkcija F ir u – duotu slenksčiu.

Seka $\{I_{\{X_i > u\}}\}$ – nepriklausomų vienodai pasiskirsčiusiais Bernulio atsitiktinių dydžių seka su pasisekimo tikimybe $p = \bar{F}(u)$. Čia pasisekimo tikimybė suprantama, kaip tikimybė, kad pavyks pakartoti nusakytą įvykį.

Taigi, pirmo pasisekimo laikas

$$L(u) = \min\{i \geq 1 : X_i > u\}, \quad (1.2.4)$$

t.y. pirmo perviršio per slenkstį u pasirodymas, yra geometrinis atsitiktinis dydis su pasiskirstymo funkcija

$$P(L(u) = k) = (1 - p)^{k-1} p, k = 1, 2, \dots \quad (1.2.5)$$

Nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai $L_1(u) = L(u)$,

$$L_{n+1}(u) = \min\{i > L_n(u) : X_i > u\}, n \geq 1,$$

seka $\{X_n\}$ apibūdina laikotarpį tarp perviršių per slenkstį u .

Įvykių $\{X_i > u\}$ pasikartojimo periodas apibūdinamas

$$EL(u) = p^{-1} = (\bar{F}(u))^{-1}, \quad (1.2.6)$$

kuris artėja į begalybę, kai $u \rightarrow \infty$.

Į visus susijusius su pasikartojimo periodu klausimus dabar gali lengvai atsakyti atitinkami geometrinio skirstinio parametrai.

Apibrėšime

$$r_k = P(L(u) \leq k) = p \sum_{i=1}^k (1-p)^{i-1} = 1 - (1-p)^k, k \in N, \quad (1.2.7)$$

čia r_k – tikimybė, kad bus bent vienas slenksčio u perviršis per k stebėjimų.

Tikimybė, kad slenkstis u bus viršijamas anksčiau, nei pasikartos įvykis:

$$P(L(u) \leq EL(u)) = P(L(u) \leq [1/p]) = 1 - (1-p)^{[1/p]},$$

čia $[x]$ reiškia sveiką x dalį.

Didelėms slenksčio u reikšmėms, t.y. kai $u \rightarrow \infty$ ir atitinkamai $p \rightarrow 0$

$$\lim_{u \uparrow \infty} P(L(u) \leq EL(u)) = \lim_{p \downarrow 0} (1 - (1-p)^{[1/p]}) = 1 - e^{-1} = 0.63212$$

Tai rodo, kad dideliems slenksčiams $L(u)$ pasikartojimo periodo vidurkis yra didesnis už jo medianą.

Pavyzdžiui, draudimo kontekste, pastatas draudžiamas tokiu pagrindu: draudimo periodas tęsis bent 50 metų ir avarinės situacijos rizika neviršija 10%. Kaip ši informacija atsiliepia pasikartojimo periodui?

$$P(L(u) \leq 50) \leq 0,1.$$

Pastato avarinė situacija i metams aprašoma įvykiu $\{X_i > u\}$,
čia X_i - kritinė komponentė, nuo kurios priklauso pastato stabilumas.

Ankstesnė sąlyga užrašoma taip:

$$P(L(u) \leq 50) = 1 - (1 - p)^{50} = 0,1.$$

Iš čia išplaukia, kad $p = 0,002105$, t.y. $EL(u) = 475$. Draudime tai reiškia “475 metų” įvykį.

Kitas svarbus klausimas yra susijęs su t metų įvykio reikalavimu slenksčio reikšmei. Pagal apibrėžimą slenksčiui u_t :

$$t = EL(u) = \frac{1}{\overline{F}(u_t)},$$

iš čia išplaukia

$$u_t = F^{\leftarrow}(1 - t^{-1}).$$

Šiame pavyzdyje, $u_{475} = F^{\leftarrow}(0,9979)$.

1.3 LAIKO EILUTĖS

1.3.1 PAGRINDINĖS SĄVOKOS

Labai dažnai kintamam reiškiniui aprašyti stebimos kokio nors kintamojo dydžio reikšmės, įgyjamos laikui bėgant ar skirtingose vietose. Pavyzdžiui, kasdiena matuojamas kritulių kiekis, fiksuojamas metinis gyventojų skaičiaus prieaugis ar daromas žmogaus kalbos įrašas. Visais atvejais turime kintančią sistemą, kurią veikia atsitiktiniai veiksniai. Jos praeitis, vienu ar kitu būdu užrašyta renkant duomenis, laikui bėgant, suteikia tam tikros informacijos apie nagrinėjamą reiškinį. Kasdieniniam kritulių kiekiui būdingi svyravimai, kurie neišreiškiami paprasta matematine formule; kalbos garsai, užrašomi fiksuojant atitinkamus įtampos svyravimus elektrinėje grandinėje, nors ir paklūsta elektromagnetizmo teorijos dėsniams, bet dėl neišvengiamų matavimo paklaidų turi atsitiktinį dėmenį.

Minėti kritulių kiekio bei gyventojų skaičiaus prieaugio stebėjimai yra diskrečiojo laiko sekų pavyzdžiai, kai matuojama kasdien, kasmet (galėtų būti ir kas sekundę). O kalbos signalo įrašas yra tolydžiojo laiko stebėjimų pavyzdys, kai rašoma tam tikrą laiką nepertraukiamai.

Visus procesus galima suskirstyti į *determinuotus*, kurių kitimą laiko bėgyje galima tiksliai aprašyti, ir *atsitiktinius*.

Tegul T yra skaičių seka arba intervalas. Visuma atsitiktinių dydžių $\{\xi_t, t \in T\}$, apibrėžtu vienoje tikimybinėje erdvėje (Ω, F, P) , vadinama *atsitiktiniu procesu*. Parametro t kitimo aibė T kartais vadinama indeksų aibe. Aibės T pavyzdžiai:

$$Z = \{0, \pm 1, \pm 2, \dots\}, N = \{0, 1, 2, \dots\}, \{-\infty, \infty\}, [0, \infty). \quad (1.3.1)$$

Realiai dažniausiai stebima viena atsitiktinio proceso realizacija (trajektorija). Jeigu fiksuosime laiko momentus t_1 ir t_2 , gausime atsitiktinius dydžius ξ_1 ir ξ_2 , kurie vadinami *proceso pjūviais* arba tiesiog atsitiktinio proceso reikšmėmis laiko momentais t_1 ir t_2 .

Kai argumento reikšmė $t = t_1$ fiksuota, proceso pjūvis ξ_1 yra atsitiktinis dydis, kurio tikimybių skirstinį nusako pasiskirstymo funkcija:

$$F_{\xi_1}(x) = P(\xi_1 \leq x), x \in \mathfrak{R}. \quad (1.3.2)$$

Vienmatė pasiskirstymo funkcija nėra viso proceso charakteristika. Ji neatspindi ryšio tarp atskirų proceso pjūvių.

Atsitiktinis procesas, kurio $T \subset Z$, dažniausiai vadinamas *laiko eilute*. Gali būti atveju, kai t yra bet koks parametras. Tačiau tradiciškai atsitiktinių sekų stebėjimai siejami su laiku ir jos vadinamos *laiko eilutėmis*.

Atsitiktinių sekų analizėje labiausiai domina ne atskiras atsitiktinis dydis ξ_t , o jų sistema $\{\xi_t\}$, priklausanti nuo parametro t , kintančio tam tikrame intervale ar įgyjančio tam tikras reikšmes. Jei tiriama sekos elgsena laiko momentais $t = 1, 2, \dots, N$, reikia nagrinėti atsitiktinių dydžių $\xi_1, \xi_2, \dots, \xi_N$ daugiamatį skirstinį.

Pasiskirstymo funkcijų rinkinys

$$\{F_{t_1, \dots, t_k}(x_1, \dots, x_k), t_1, \dots, t_k \in T, k = 1, 2, \dots\} \quad (1.3.3)$$

vadinamas *proceso ξ daugiamatiais skirstiniais*, jei $F_{t_1, \dots, t_k}(\cdot)$ yra atsitiktinių vektorių $(\xi_{t_1}, \dots, \xi_{t_k})$ pasiskirstymo funkcijos, t. y.

$$F_{t_1, \dots, t_k}(x_1, \dots, x_k) = P\{\xi_{t_1} < x_1, \dots, \xi_{t_k} < x_k\}. \quad (1.3.4)$$

Tik iš daugiamatės pasiskirstymo funkcijos bei jos charakteristikų galima spręsti apie tiriamosios sekos savybes. Praktiniuose uždaviniuose dažniausiai neturima tokios išsamios informacijos ir apsiribojama prielaidomis, daugiau ar mažiau adekvačiomis tikrovei.

Vienas pirmųjų darbų iš laiko eilučių yra Yule (1927) straipsnis, kuriame pirmą kartą buvo pasiūlytas autoregresijos modelis Saulės aktyvumo duomenims aprašyti. Darbai laiko eilučių srityje suaktyvėjo apie 1955 metus, pradėjus naudoti kompiuterius, o nuo 1970 metų tapo viena greičiausiai besiplėtojančių tyrimo sričių atsitiktinių procesų ir matematinės statistikos sankirtoje. Kompiuteriai atvėrė vis daugiau galimybių, buvo sukurta daug metodų, leidžiančių sukurti gerus statistinius

modelius įvairiems duomenims. Daugelyje statistinės analizės paketų laiko eilučių analizės procedūros yra išskirtos į atskirą komponentę. Pavyzdžiui, sistemoje SPSS yra komponentė SPSS / Trends, sistemoje SAS – SAS / ETS, kurios skirtos laiko eilučių analizei.

1.3.2 STATISTINĖ ATSTITIKTINIŲ SEKŲ ANALIZĖ IR JŲ MODELIAI

Skiriamasis statistinės laiko eilučių analizės bruožas yra nagrinėjamų atsitiktinių dydžių *statistinių ryšių* ir *koreliacijos* įvertinimas: tariama, kad kiekvienas atsitiktinis dydis ξ_t yra priklausomas nuo buvusių sekos reikšmių $\xi_{t-1}, \xi_{t-2}, \dots$, o kartais ir nuo busimų reikšmių $\xi_{t+1}, \xi_{t+2}, \dots$.

Statistinė analizė turi atskleisti šio priklausomumo pobūdį, dėsningumus ir sukurti statistinį modelį, tinkantį ne tik nagrinėjamiems, bet ir kitiems tos pačios kilmės reiškiniams aprašyti.

Pati paprasčiausia seka, neturinti jokių vidinių statistinių ryšių, yra *nepriklausomų* atsitiktinių dydžių seka $\{\varepsilon_t\}$. Tegul bet kurio atsitiktinio dydžio ε_t vidurkis μ ir dispersija σ^2 nepriklauso nuo laiko:

$$E\{\varepsilon_t\} = \mu, \quad E\{(\varepsilon_t - \mu)^2\} = \text{var}(\varepsilon_t) = \sigma^2, \quad \forall t, \quad (1.3.5)$$

tai bet kurie du dydžiai $\varepsilon_t, \varepsilon_s, \forall t \neq s$ yra *nekoreliuoti*:

$$\text{cov}(\varepsilon_t, \varepsilon_s) = E[(\varepsilon_s - \mu)(\varepsilon_t - \mu)] = E(\varepsilon_s - \mu)E(\varepsilon_t - \mu) = 0$$

Bendruoju atveju seka $\{\xi_t\}$ turi sudėtingą struktūrą. Jos vidurkis $\mu_t = E\{\xi_t\}$ gali būti laiko funkcija ir kisti laikui bėgant. Be to, laiko eilutės svyravimai apie vidurkį

$$\xi_t^0 = \xi_t - \mu_t \quad \text{ir} \quad \xi_s^0 = \xi_s - \mu_s \quad (1.3.6)$$

gali būti koreliuoti dydžiai. Kuriant modelį būtina atsižvelgti į šias sekos savybes.

Paprastumo dėlei tarkime, kad nagrinėjama ξ_t^0 tipo seka ir visiem t turime $E\{\xi_t^0\} = 0$. Dabar belieka įvertinti dydžių tarpusavio priklausomumą, atskleisti statistinius ryšius tarp $\dots, \xi_{t-2}^0, \xi_{t-1}^0, \xi_t^0, \xi_{t+1}^0, \xi_{t+2}^0, \dots$ ir surasti koku būdu $\{\xi_t^0\}$ gali būti keičiama nepriklausomųjų dydžių seka $\{\varepsilon_t\}$. Taigi bendriausia prasme, laiko eilutės modelis būtų išreiškiamas lygtimi

$$\varphi\{\dots, \xi_{t-2}^0, \xi_{t-1}^0, \xi_t^0, \xi_{t+1}^0, \xi_{t+2}^0, \dots\} = \varepsilon_t, \quad (1.3.7)$$

čia $\varphi(\bullet)$ - tam tikra funkcija.

Nepriklausomų atsitiktinių dydžių ε_t seka dešinėje (1.3.7) lygties pusėje reiškia, kad modelis paaiškino visus ryšius laiko eilutėje $\{\xi_t^0\}$ ir jį pritaikius liko paprasčiausia struktūra – nepriklausomi dydžiai.

Bendras laiko eilučių teorijos uždavinys gali būti formuluojamas taip: tarkime turime laiko eilutės $\{\xi_t\}$ stebėjimus $\xi_1, \xi_2, \dots, \xi_N$, reikia surasti tokią funkciją $\varphi(\bullet)$, kuri keičia $\{\xi_t\}$ seką į nepriklausomųjų dydžių seką $\{\varepsilon_t\}$.

Jeigu funkcija $\varphi(\bullet)$ yra tokia, kad (1.3.7) lygtis gali būti išspręsta ξ_t atžvilgiu:

$$\xi_t = f\{\dots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t, \varepsilon_{t+1}, \varepsilon_{t+2}, \dots\}, \quad (1.3.8)$$

čia $f(\bullet) = \varphi^{-1}(\bullet)$ yra funkcija, atvirkštinė funkcijai $\varphi(\bullet)$, tai reiškia, kad turime geriausią modelį sekai $\{\xi_t\}$.

Suprantama, kad kai funkcija $\varphi(\bullet)$ yra bet kuri iš visų galimų funkcijų klasės, o stebėjimų skaičius N yra baigtinis, šis uždavinys neišsprendžiamas. Tenka apriboti nagrinėjamų funkcijų klasę, pavyzdžiui, kintamųjų $\dots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t, \varepsilon_{t+1}, \varepsilon_{t+2}, \dots$ tiesinių funkcijų klase ir tada spręsti minėtą uždavinį.

1.3.3 STACIONARIŲ SEKŲ SĄVOKOS

Negriežtai šnekant, $\{\xi_t\}$ seka vadinama *stacionariąja*, jeigu jos savybės nekinta laikui bėgant. Griežtai formuluojant, skiriamas stacionarumas plačiąja ir siaurąja prasme.

Procesas ξ_t vadinamas *stacionariu siaurąja prasme*, jei jo daugiamačiai pasiskirstymai nepriklauso nuo postūmio laike, t. y. :

$$t_1, \dots, t_k \in T, k = 1, 2, \dots \quad F_{t_1, \dots, t_k}(\bullet) = F_{t_1 + \tau, \dots, t_k + \tau}(\bullet), \quad t_i + \tau \in T \quad (1.3.9)$$

Laiko eilučių analizėje dažniausiai naudojamas kitas apibrėžimas.

Procesas ξ_t vadinamas *stacionariu plačiąja prasme*, jei jo matematinis vidurkis ir kovariacinė funkcija nepriklauso nuo poslinkio laike, t. y. jei

$$\forall t, s \in T \quad \mu(t) = \mu(0), \quad (1.3.10)$$

$$R(t, s) = R(t - s, 0). \quad (1.3.11)$$

Akivaizdu, kad jei procesas ξ_t yra stacionarus siaurąja prasme ir turi dispersiją, tai jis yra ir stacionarus plačiąja prasme, bet ne atvirkščiai. Tačiau, jei procesas ξ_t yra Gauso (jei jo daugiamačiai pasiskirstymai yra Gauso pasiskirstymo funkcijų rinkinys), jam abu apibrėžimai sutampa.

Taigi stacionaraus proceso ξ_t matematinis vidurkis nekinta laike $E\xi_t = \mu$, o kovariacinė funkcija yra vieno argumento funkcija:

$$\text{cov}(\xi_{t+\tau}, \xi_t) = E[(\xi_{t+\tau} - \mu)(\xi_t - \mu)] = R(\tau), \quad \forall t \in T. \quad (1.3.12)$$

Ši funkcija yra neneigiamai apibrėžta:

$$\forall t_1, \dots, t_k \in T, x_1, \dots, x_k : \sum_{i,j=1}^k R(t_i - t_j) x_i x_j \geq 0. \quad (1.3.13)$$

Akivaizdu, kad $R(\bullet)$ yra lyginė funkcija, $R(-\tau) = R(\tau)$.

Kiekvienam τ funkciją $R(\tau)$ padalijus iš $R(0) \equiv \sigma^2$, turime *koreliacinę funkciją*

$$r(\tau) = \text{cor}(\xi_{t+\tau}, \xi_t) = \frac{R(\tau)}{R(0)}. \quad (1.3.14)$$

Ši funkcija taip pat lyginė, neneigiamai apibrėžta ir $|r(\tau)| \leq 1$. Reikšmė $r(\tau)$ parodo kiek stipriai proceso reikšmės dabartyje tiesiškai priklauso nuo reikšmės prieš τ laiko vienetų. Koreliacinių ryšių žinojimas palengvina laiko eilučių modelio parinkimą ir identifikavimą.

1.3.4 ATSITIKTINIŲ SEKŲ MODELIAI, JŲ OPERATORINĖ FORMA

Kaip jau buvo paminėta, bendrasis statistinių modelių kūrimo uždavinys išsprendžiamas funkcijų $H(\bullet)$ klasę susiaurinus iki tiesinių funkcijų klasės. Todėl nagrinėjamas toks tiesinis modelis:

$$\sum_{u=-\infty}^{\infty} h_u \xi_{t-u} = \varepsilon_t, \quad (1.3.15)$$

čia $\{h_u\}$ - konstantų seka.

(1.3.15) lygtis išreiškia bendrąjį tiesinį modelį, kuriame ξ_t gali priklausyti ir nuo buvusiųjų reikšmių $\xi_{t-1}, \xi_{t-2}, \dots$, ir nuo būsimų reikšmių $\xi_{t+1}, \xi_{t+2}, \dots$. Tačiau dažniausiai yra tariama, kad $h_u = 0$, kai $u < 0$ ir nagrinėjamas vienpusis modelis

$$\sum_{u=0}^{\infty} h_u \xi_{t-u} = \varepsilon_t. \quad (1.3.16)$$

(1.3.16) lygtyje ξ_t priklauso nuo ε_t ir jo buvusiųjų reikšmių $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$, bet negali priklausyti nuo būsimų reikšmių $\varepsilon_{t+1}, \varepsilon_{t+2}, \dots$.

Apibrėžkime *poslinkio atgal operatorių* L :

$$L^k \xi_t = \xi_{t-k}. \quad (1.3.17)$$

Pažymėjus

$$H(z) = \sum_{u=0}^{\infty} h_u z^u, \quad (1.3.18)$$

modelį (1.3.16) galima užrašyti operatorine forma:

$$H(L)\xi_t = \varepsilon_t, \quad (1.3.19)$$

ir formaliai išspręst ξ_t atžvilgiu:

$$\xi_t = H^{-1}(L)\varepsilon_t, \quad (1.3.20)$$

jei tik $H(z)$ neturi nulių vienetiniame skritulyje ir ant jo kontūro: $|z| \leq 1$. Funkciją $H^{-1}(z)$ užrašius skleidiniu $H^{-1}(z) = g_0 + g_1z + g_2z^2 + \dots$, (1.3.20) išraiška lygiavertė

$$\xi_t = \sum_{u=0}^{\infty} g_u \varepsilon_{t-u} \quad (1.3.21)$$

arba
$$\xi_t = G(L)\varepsilon_t, \quad (1.3.22)$$

čia
$$G(z) = \sum_{u=0}^{\infty} g_u z^u.$$

Bendrą tiesinio modelio sudarymo uždavinį dabar galima formuluoti taip: turimiems $\{\xi_t\}$ sekos stebėjimams $\xi_1, \xi_2, \dots, \xi_N$ įvertinti tokią nežinomų parametru seką $\{h_u\}$, kad (1.3.16) pakeistų $\{\xi_t\}$ seką į nepriklausomųjų dydžių ε_t seką.

Reikalaujama iš baigtinio skaičiaus duomenų įvertinti begalinį skaičių parametru, todėl bendruoju atveju tai vėl neišsprendžiama problema. Privaloma toliau siaurinti klasę, kurioje ieškotume sprendinio. Priartėjame prie baigtinio parametru skaičiaus tiesinių modelių klasės, kurioje anksčiau formuluotas uždavinys jau išsprendžiamas. Tam yra naudojami ARMA modeliai, kurie aprašo stacionarų procesą kaip balto triukšmo tiesinį filtrą su baigtiniu skaičiumi parametru. Autoregresijos ir slenkančiojo vidurkio procesai yra ARMA proceso atskiri atvejai.

1.3.5 STACIONARŪS TIESINIAI MODELIAI

Stacionarus procesas ξ_t vadinamas *p* eilės autoregresijos procesu (*AR(p)*), jei jis išreiškiamas:

$$\xi_t = \mu + a_1 \xi_{t-1} + \dots + a_p \xi_{t-p} + \varepsilon_t, \quad t \in Z, \quad (1.3.23)$$

čia ε_t - baltas triukšmas.

Pažymėję $P(z) = 1 - a_1z - \dots - a_pz^p$, lygybę (1.3.23) galima užrašyti:

$$P(L)\widehat{\xi}_t = \varepsilon_t, \quad \widehat{\xi}_t = \xi_t - E\xi_t. \quad (1.3.24)$$

Be to, $\mu = P(1)E\xi_t$.

Stacionarus procesas ξ_t vadinamas *q* eilės slenkančiojo vidurkio (angl. – moving average) procesu (*MA(q)*), jei jis išreiškiamas:

$$\xi_t = \mu + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}, \quad t \in Z, \quad (1.3.25)$$

čia ε_t - balto triukšmo procesas.

Pažymėję $Q(z) = 1 + b_1 z + \dots + b_q z^q$, gauname $\mu = E\xi_t$ ir

$$\widehat{\xi}_t = Q(L)\varepsilon_t. \quad (1.3.26)$$

MA procesas yra reguliarus procesas, kuris gaunamas tiesiškai nufiltruojant baltą triukšmą su baigtiniu skaičiumi nenulinių koeficientų, t.y. filtro dažnuminė charakteristika yra trigonometrinis polinomas.

Stacionarus procesas ξ_t vadinamas *ARMA(p,q)* procesu, jei

$$\xi_t = \mu + a_1 \xi_{t-1} + \dots + a_p \xi_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}, \quad t \in Z, \quad (1.3.27)$$

čia ε_t - balto triukšmo procesas.

Stacionarus procesas ξ_t , tenkinantis lygybę (1.3.27), egzistuoja tada ir tik tada, jei polinomas $P(z)$ neįgauna nulinių reikšmių ant vienetinio apskritimo (kompleksinių skaičių plokštumoje):

$$P(z) \neq 0, \quad |z| = 1.$$

1.3.6 NESTACIONARŪS TIESINIAI MODELIAI

Tarp ekonominių kintamųjų nėra sunku rasti tokius, kurie akivaizdžiai nėra stacionarūs – santaupas, šalies BVP ir pan., kadangi šie rodikliai normaliom sąlygomis auga. Tačiau jų pokyčiai (pvz. per metus) dažnai elgiasi kaip stacionarūs procesai.

Atsitiktinis procesas ξ_t vadinamas *d – eilės integruotu* (žymima $\xi \in I(d)$), jei jo *d* eilės pokyčiai yra stacionarus procesas, o *d – 1* eilės pokyčiai nėra stacionarūs.

Pažymėkime:

$$\Delta \xi_t = \xi_t - \xi_{t-1} = (1 - L)\xi_t.$$

Tada: $\Delta^2 \xi_t = \Delta \xi_t - \Delta \xi_{t-1} = (1 - L)^2 \xi_t = \xi_t - 2\xi_{t-1} + \xi_{t-2}.$

Bendru atveju:

$$\Delta^d \xi_t = \Delta^{d-1} \xi_t - \Delta^{d-1} \xi_{t-1} = (1 - L)^d \xi_t. \quad (1.3.28)$$

Taigi $\xi \in I(d) \Rightarrow \eta_t = \Delta^d \xi_t$ yra stacionarus procesas. Jei pats ξ_t yra stacionarus, žymima $\xi \in I(0)$.

Tuo atveju, kai $\xi \in I(1)$, galioja lygybė

$$\xi_t = \xi_0 + \sum_{\tau=1}^t \eta_\tau, \quad t = 1, 2, \dots \quad (1.3.29)$$

čia η_τ - stacionarus procesas. Taigi ξ_t gaunamas sumuojant stacionaraus proceso reikšmes. Iš čia ir pavadinimas – integruoti procesai.

Atsitiktinis procesas $\xi_t \in I(d)$ vadinamas *ARIMA(p,d,q)* procesu, jei jo *d* eilės pokyčiai $\eta_t = (1 - L)^d \xi_t$ yra *ARMA(p,q)* stacionarus procesas. Taigi, galioja lygybė:

$$P(L) \cdot (1-L)^d \widehat{\xi}_t = Q(L)\varepsilon_t, \quad (1.3.30)$$

čia $P(z)$ ir $Q(z)$ - p ir q eilės polinomai atitinkamai, o ε_t - balto triukšmo procesas.

1.3.7 BENDRA TIESINIO PROGNOZAVIMO TEORIJA IR STATISTINŲ MODELIŲ PANAUDOJIMAS PROGNOZEI

Statistinio modelio sukūrimas nagrinėjamiems duomenims nėra savitiksliis uždavinys. Kiekvienas modelis yra tam tikra tikrovės idealizacija, todėl galimą modelį panaudot sprendžiant tokius uždavinius:

1. prognozuoti būsimas sekos reikšmes;
2. modeliuoti daugiau panašių realizacijų;
3. atkurti trūkstamas reikšmes stebėjimų sekoje;
4. išgryninti stebėjimus, atmetant reikšmes, atsiradusias sekoje dėl pašalinio poveikio.

Prognozė suprantama kaip būsimų proceso reikšmių įvertinimas remiantis turimomis proceso reikšmėmis.

Tarkime, stebime atsitiktinį vektorių $X = (X_1, \dots, X_n)^T$. Atsitiktinio dydžio Y tiesinė prognozė:

$$\hat{Y} = a + b_1 X_1 + \dots + b_n X_n = a + b^T X \quad (1.3.31)$$

Prognozės tikslumo matas – vidutinė kvadratinė paklaida:

$$\Delta = \Delta(a, b) = E\varepsilon^2, \quad \varepsilon = Y - \hat{Y} \quad (1.3.32)$$

Vidutinė kvadratinė paklaida gaunama mažiausia, kai koeficientai parenkami taip, kad

$$E\varepsilon = 0, \quad \text{cov}(\varepsilon, X) = 0$$

Optimalūs koeficientai randami iš lygčių

$$\text{cov}(\varepsilon, X) = \text{cov}(Y, X) - \text{cov}(b^T X, X) = R_{YX} - b^T R_{XX} = 0$$

$$E\varepsilon = EY - a - b^T EX = 0$$

Jei kovariacinė matrica R_{XX} neišsigimusi, tai sprendinys vienas:

$$b^* = R_{XX}^{-1} R_{XY}, \quad a^* = EY - b^{*T} EX.$$

1.3.7.1 STACIONARUS TIESINIS PROCESAS

Tarkime, ξ_t yra stacionarus procesas su vidurkiu $E\xi_t = \mu$ ir kovariacine funkcija $R(\tau) = \text{cov}(\xi_t, \xi_{t+\tau})$.

Stebima imtis (ξ_1, \dots, ξ_n) , o reikia prognozuoti ξ_s , kur $s > n$. Optimali tiesinė prognozė

$$\hat{\xi}_s = \alpha + \beta_1 \xi_1 + \dots + \beta_n \xi_n$$

gaunama, kai:

$$\beta = R_{XX}^{-1} R_{X\xi_s}, \quad \alpha = \mu - \beta^T EX = \mu(1 - \beta_1 - \dots - \beta_n),$$

čia $X = (\xi_1, \dots, \xi_n)^T$.

Todėl $R_{XX} = [R(i-j)]_{i=1, \dots, n}^{j=1, \dots, n}$, $R_{X\xi_s} = (R(s-1), R(s-2), \dots, R(s-n))^T$.

1.3.7.2 NESTACIONARUS TIESINIS PROCESAS

Jei stebime $\xi_0, \xi_1, \dots, \xi_n$, kur ξ_t yra ARIMA (p, d, q) procesas, tai pažymėję

$$\eta_t = (1-L)^d \xi_t$$

randame $\eta_d, \eta_{d+1}, \dots, \eta_n$. Iš šios imties įvertiname nežinomus polinomų $P(z)$ ir $Q(z)$ koeficientus, apskaičiuojame proceso η_t koreliacinę funkciją $r(\tau)$ ir η taikome bendrą tiesinio prognozavimo teoriją.

Gavus prognozes $\hat{\eta}_t, t = n+1, n+2, \dots$, atitinkamas prognozes $\hat{\xi}_t$ galima rasti iš išraiškos:

$$\eta_t = (1-L)^d \xi_t = \sum_{k=0}^d (-1)^k \binom{d}{k} \xi_{t-k},$$

čia $\binom{d}{k} = \frac{d!}{k!(d-k)!}$.

Žinodami η_t ir ξ_1, \dots, ξ_n reikšmes ξ_t , rekurentiniu būdu galime surasti

$$\xi_t = \eta_t - \sum_{k=1}^d (-1)^k \binom{d}{k} \xi_{t-k}, \quad t = n+1, n+2, \dots$$

Nustatant tiriamo proceso ξ_t integruotumo eilę, praktikoje dažniausiai naudojamas *Dickey* – *Fuller* testas arba jo modifikacijos. Šis testas naudoja AR(1) modelį

$$\xi_t = \mu + a\xi_{t-1} + \varepsilon_t$$

ir tikrina hipotezę $H_0: a = 1$, esant alternatyviai hipotezei: $H_1: a < 1$. Jei hipotezė H_0 priimama, daroma išvada, kad ξ_t nėra stacionarus, ir pereinama prie analogiško $\Delta\xi_t$ nagrinėjimo. Jei ir $\Delta\xi_t$ nėra stacionarus, tiriami antros eilės pokyčiai $\Delta^2\xi_t$ ir taip toliau.

1.3.8 MODELIO EILĖS NUSTATYMO BŪDAI

1.3.8.1 MAŽIAUSIOS KANONINĖS KORELIACIJOS METODAS (SCAN)

Šis metodas padeda preliminariai nustatyti stacionaraus arba nestacionaraus ARMA modelio eilę. Tarkime, yra stacionari arba nestacionari laiko eilutė $\{z_t : 1 \leq t \leq n\}$, turinti pataisytą vidurki $\tilde{z}_t = z_t - \mu$, autoregresijos eilę $p+d$ ir slenkamojo vidurkio eilę q .

Nežinomai autoregresijos eilei $m = p_{\min}, \dots, p_{\max}$ ir nežinomai slenkamojo vidurkio eilei $j=q_{\min}, \dots, q_{\max}$ nustatyti atliekami tokie veiksmai:

1) Tarkime $Y_{m,t} = (\tilde{z}_t, \tilde{z}_{t-1}, \dots, \tilde{z}_{t-m})'$. Apskaičiuojama $(m+1) \times (m+1)$ formato matrica:

$$\hat{\beta}(m, j+1) = \left(\sum_t Y_{m,t-j-1} Y'_{m,t-j-1} \right)^{-1} \left(\sum_t Y_{m,t-j-1} Y'_{m,t} \right)$$

$$\hat{\beta}^*(m, j+1) = \left(\sum_t Y_{m,t} Y'_{m,t-j-1} \right)^{-1} \left(\sum_t Y_{m,t} Y'_{m,t-j-1} \right), \quad t = j+m+2, \dots, n$$

$$\hat{A}^*(m, j) = \hat{\beta}^*(m, j+1) \hat{\beta}(m, j+1)$$

2) Iš $\hat{A}^*(m, j)$ randama mažiausia tikrinė reikšmė $\hat{\lambda}^*(m, j)$ ir atitinkamas normalizuotas tikrinis vektorius $\Phi_{m,j} = (1, -\phi_1^{(m,j)}, -\phi_2^{(m,j)}, \dots, -\phi_m^{(m,j)})$.

$\hat{\lambda}^*(m, j)$ yra kvadratinis kanoninis koreliacijos įvertis.

3) Naudojant $\Phi_{m,j}$ kaip proceso AR(m) koeficientus, pagal formulę

$$w_t^{(m,j)} = \tilde{z}_t - \phi_1^{(m,j)} \tilde{z}_{t-1} - \phi_2^{(m,j)} \tilde{z}_{t-2} - \dots - \phi_m^{(m,j)} \tilde{z}_{t-m}$$

įvertinamos liekanos, kai $t = j+m+1, \dots, n$.

4) Iš liekanų autokoreliacijos $r_k(w)$, pagal formulę

$$\text{var}(\hat{\lambda}^*(m, j)^{1/2}) \approx \frac{d(m, j)}{n - m - j},$$

čia $d(m, j) = \left(1 + 2 \sum_{i=1}^{j-1} r_i(w^{(m,j)}) \right)$, įvertinama kvadratinųjų kanoninių autokorelacijų standartinė paklaida.

Kaip atpažinimo kriterijus, naudojama statistika $c(m, j) = -(n - m - j) \ln(1 - \hat{\lambda}^*(m, j) / d(m, j))$, kuri asimptotiškai yra χ_1^2 , jei $m = p+d$ ir $j \geq q$, arba jeigu $m \geq p+d$ ir $j = q$. Kai $m > p$ ir $j < q$, yra daugiau nei viena teorinė nulinė kanoninė koreliacija tarp $Y_{m,t}$ ir $Y_{m,t-j-1}$. Kadangi $\hat{\lambda}^*(m, j)$ yra

mažiausia kanoninė koreliacija kiekvienam (m,j) , $c(m,j)$ procentiliai yra mažesni už χ_1^2 procentilius. Todėl galima priimti χ_1^2 . Kai $m < p$ ir $j < q$, jokios išvados apie $c(m,j)$ padaryti negalima.

SCAN metodo lentelė, sudaryta iš $c(m,j)$, padeda nustatyti kuri iš tikrinių reikšmių $\hat{\lambda}^*(m,j)$ reikšmingai skiriasi nuo nulio.

1.3.8.2 IŠPLĖSTOSIOS IMTIES AUTOKORELIACINĖS FUNKCIJOS METODAS (ESACF)

Šis metodas padeda preliminariai nustatyti stacionaraus arba nestacionaraus ARMA modelio eilę. Tarkime, yra stacionari arba nestacionari laiko eilutė $\{z_t : 1 \leq t \leq n\}$, turinti pataisytą vidurki $\tilde{z}_t = z_t - \mu$, autoregresijos eilę $p+d$ ir slenkamojo vidurkio eilę q . ESACF metodas analizuoja autokoreliacines funkcijas susietas su filtruotomis sekomis, turinčiomis pavidalą

$$w_t^{(m,j)} = \hat{\Phi}_{(m,j)}(B)\tilde{z}_t = \tilde{z}_t - \sum_{i=1}^m \hat{\phi}_i^{(m,j)}\tilde{z}_{t-i},$$

čia B – postūmio atgal operatorius,

$m = p_{\min}, \dots, p_{\max}$ – nežinoma autoregresijos eilė,

$j = q_{\min} + 1, \dots, q_{\max} + 1$ – slenkamojo vidurkio nežinoma eilė,

$\hat{\phi}_i^{(m,j)}$ - autoregresijos parametrų įverčiai, tariant, kad seka yra ARMA(m,j) procesas.

Grynai autoregresiniams modeliams ($j=0$) $\hat{\phi}_i^{(m,0)}$ įvertinimui naudojamas įprastas mažiausių kvadratų metodas. ARMA modeliams įverčiai apskaičiuojami pagal rekurentinę formulę, pradedant nuo $\hat{\phi}_i^{(m,0)}$:

$$\hat{\phi}_i^{(m,j)} = \hat{\phi}_i^{(m+1,j-1)} - \hat{\phi}_{i-1}^{(m,j-1)} \frac{\hat{\phi}_{m+1}^{(m+1,j-1)}}{\hat{\phi}_m^{(m,j-1)}}$$

Filtruotų sekų imties autokoreliacijos funkcijos j -asis vėlinimas – $w^{(m,j)}_t$, yra išplėtosios imties autokoreliacijos funkcija ir žymima $r_{j(m)} = r_j(w^{(m,j)})$. Standartinės $r_{j(m)}$ paklaidos apskaičiuojamos naudojant imties autokoreliacinės funkcijos dispersijos aproksimaciją:

$$\text{var}(r_{j(m)}) \approx (1 + \sum_{i=1}^{j-1} r^2(w^{(m,j)}))$$

Jeigu tikrasis modelis yra ARMA($p+d,q$), tai filtruotos sekos $w^{(m,j)}_t$, išplaukia iš MA(q), kai $j \geq q$ modelio:

$$r_{j(p+d)} \approx 0, j > q$$

$$r_{j(p+d)} \neq 0, j = q$$

Išplėstosios imties autokoreliacija tenkina

$$r_{j(m)} \approx 0, j - q > m - p - d \leq 0,$$

$$r_{j(m)} \neq c(m - p - d, j - q), 0 \leq j - q \leq m - p - d,$$

čia $c(m - p - d, j - q)$ konstanta nelygi nuliui.

Modelio ARMA eilė nustatoma iš ESACF metodo lentelės, kuri sudaroma iš $r_{j(m)}$ eilėms $m = p_{\min}, \dots, p_{\max}$ ir $j = q_{\min} + 1, \dots, q_{\max} + 1$.

2. TIRIAMOJI DALIS

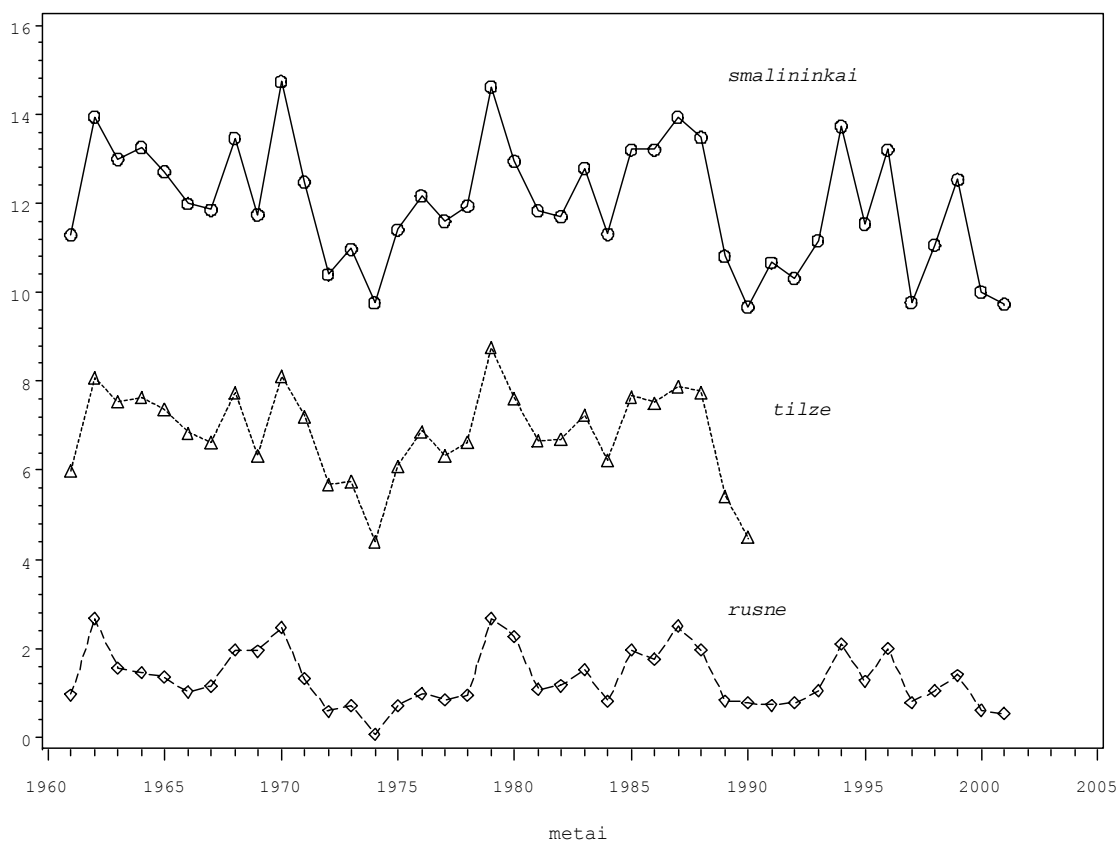
2.1 KINTAMŪJŲ APIBŪDINIMAS IR BENDROJI ANALIZĖ

Nagrinėjami hidrologiniai stebėjimai – Nemuno žemupio deltoje Smalininkų, Tilžės bei Rusnės matavimo postuose užfiksuoti vandens lygiai. Rusnės bei Smalininkų matavimo postuose užfiksuotų vandens lygių imtis sudaro 1947 – 2001 metų stebėjimai. Tilžės matavimo posto vandens lygių imtį sudaro 1947 – 1991 metų stebėjimai. Taip yra dėl to, kad Lietuvai atgavus nepriklausomybę šis matavimo postas liko Kaliningrado srityje.

Priklausomai nuo to, kokiams vandens poste buvo fiksuojami vandens lygiai, kintamieji atitinkamai pavadinti *rusne*, *tilze* ir *smalininkai*.

Turimas stebėjimų imtis būtų tikslinga padalinti į dvi dalis: iki Kauno hidroelektrinės paleidimo 1960 metais ir po jo. Kauno hidroelektrinei pradėjus veikti visu pajėgumu, didžiausi metiniai vandens lygiai Nemuno žemupio deltoje sumažėjo. Todėl atliekamoje analizėje bus nagrinėjami stebėjimai, kurie buvo užfiksuoti po 1960 metų.

vandens lygis, m



2.1 pav. Kintamųjų stebėjimų imtys

Duomenų sklaidos charakteriai yra panašūs, skiriasi tik savo dydžiu. Taip yra dėl skirtingų nulinio altitudžių šiuose matavimo postuose: Smalininkų – 7,38 m; Tilžės – 2,06 m; Rusnės – 1,53 m.

Imčių statistikos parodytos 2.1 lentelėje:

2.1 lentelė

Analizuojamų kintamųjų imčių statistikos

Statistika	Kintamasis		
	<i>rusne</i>	<i>tilze</i>	<i>smalininkai</i>
Imties stebėjimų skaičius	41	30	41
Mažiausia reikšmė	0,08	4,39	9,67
Vidurkis	1,324	6,839	11,996
Dispersija	0,420	1,082	1,904
Pirmas kvartilis	0,81	6,23	11,06
Mediana	1,15	6,87	11,85
Trečias kvartilis	1,94	7,64	13,20
Didžiausia reikšmė	2,66	8,76	14,73

2.2 EKSTREMALIŲJŲ ĮVYKIŲ STATISTINIŲ METODŲ VANDENS LYGIAMS TAIKYMAS

2.2.1 REKORDINIAI MOMENTAI

Šiuo metodu tikrinama, ar analizuojamas imtis sudaro nepriklausomi vienodai pasiskirstę dydžiai. Kiekvienoje imtyje pirmasis stebėjimas X_1 apibrėžiamas kaip rekordinis momentas.

Tarkime, kad dydžiai imtyse *rusne*, *tilze* ir *smalininkai* nepriklausomi ir vienodai pasiskirstę. Tuomet teorinis rekordinių momentų skaičius 41 stebėjimams yra

$$EN_{41} = 1 + \sum_{k=2}^{41} P(X_k > M_{k-1}) = 4,3029$$

Teorinis rekordinių stebėjimų skaičius 31 stebėjimams:

$$EN_{31} = 1 + \sum_{k=2}^{31} P(X_k > M_{k-1}) = 4,0272$$

Rekordinių momentų skaičius kintamųjų *rusne* ir *smalininkai* imtyse lygus 3. Kintamojo *tilze* imtyje rekordinių momentų skaičius lygus 4.

Turi būti tenkinamos sąlygos: $EN_{41} - \ln 41 \rightarrow \gamma$ ir $EN_{31} - \ln 31 \rightarrow \gamma$,

čia $\gamma = 0,5772$ - Eulerio konstanta.

Kintamajam *tilze*:

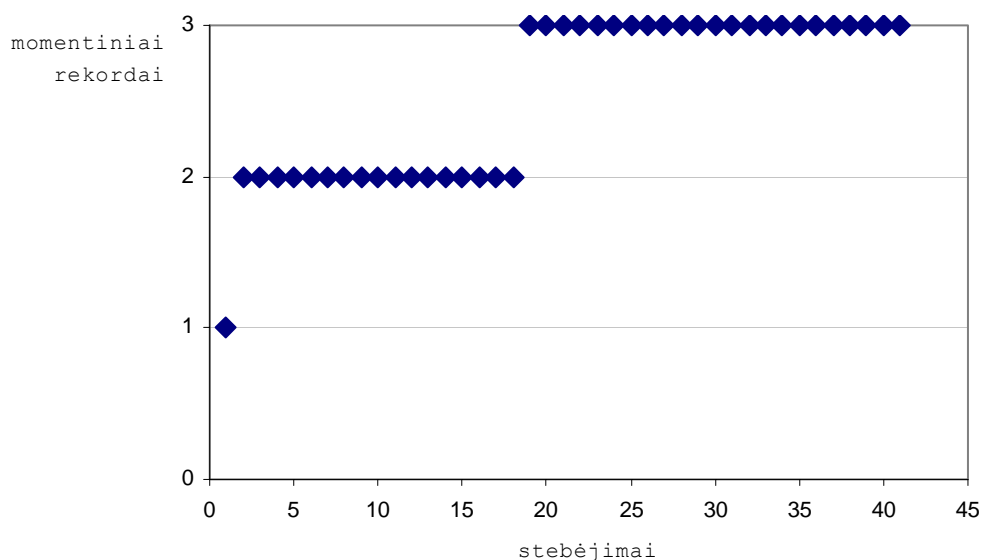
$$4,0272 - 3,4340 = 0,594 \approx 0,5932$$

Kintamiesiems *rusne* ir *smalininkai*:

$$4,3029 - 3,7136 = 0,5893 \approx 0,5772$$

Taigi, nėra pagrindo atmesti hipotezės, kad duomenys yra nepriklausomi ir vienodai pasiskirstę.

Kintamojo *rusne* rekordiniai momentai parodyti 2.2 pav. Kintamųjų *tilze* ir *rusne* rekordiniai momentai pateikti prieduose.



2.2 pav. Kintamojo *rusne* rekordiniai momentai

2.2.2 PASIKARTOJIMO PERIODAS

Nagrinėjami duomenys – didžiausios potvynių metu užtvindytos teritorijos (Šilutės rajono civilinės ir priešgaisrinės saugos tarnybos duomenimis), kurių suvestinė parodyta 2.2 lentelėje.

2.2 lentelė

Didžiausių Šilutės rajono potvynių suvestinė	
Metai	Užtvindyta teritorija, ha
1975	48 000
1978	47 000
1982	45 000
1994	48 000
1999	45 000
2002	23 000

1975 ir 1994 metais potvynių metu buvo užtvindytos vienodai didelės teritorijos: 48 000 ha. Laiko tarpas tarp šių įvykių lygus 19 metų. Taigi galima apibrėžti “19 metų” įvykį – “ypatingai didelį potvynį” taip: potvynis, kurio metu užtvindoma 48 000 ha teritorija.

$$EL(u) = \frac{1}{p} = 19, \text{ todėl "pasisekimo" tikimybė}$$

$$p = \frac{1}{19} = 0,05263.$$

Čia "pasisekimo" tikimybė suprantama kaip tikimybė, kad pasiseks pakartoti nusakytą įvykį.

Tokiu būdu, tikimybė per k metų bent vieną kartą įvykti "ypatingai dideliame potvyniui", t.y. tokiam, kurio metu užtvindyta teritorija sudaro 48 000 ha, apskaičiuojama:

$$r_k = p \sum_{i=1}^k (1-p)^{i-1} = 1 - (1-p)^k = 1 - (1-0.05263)^k, \quad k \in \mathbf{N}.$$

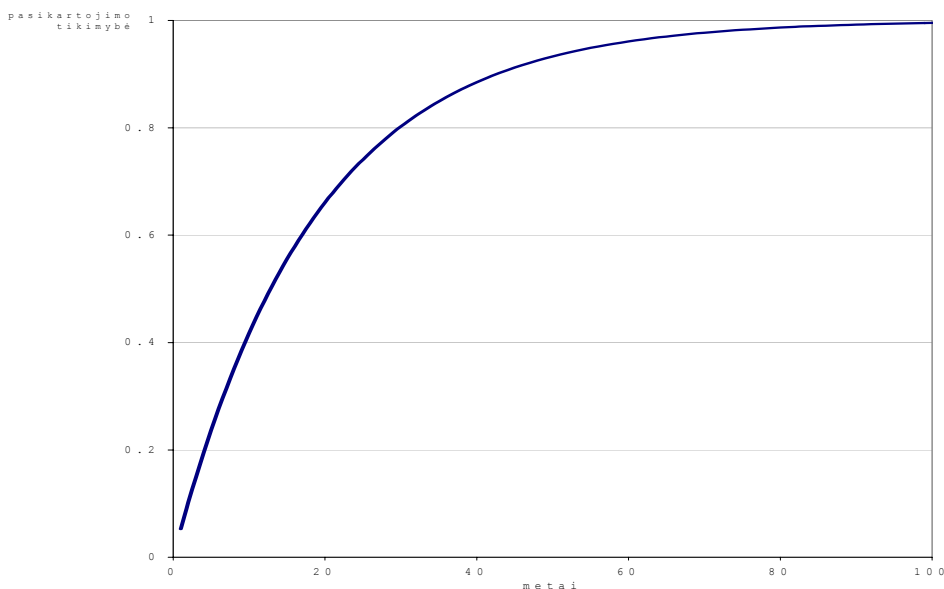
Apskaičiuotos tikimybės pavaizduotos 2.3 lentelėje:

2.3 lentelė

Tikimybė bent vieną kartą pasirodyti įvykiui "ypatingai didelis potvynis"
 k metų laikotarpyje

k	r_k
1	0,05263
2	0,10249
3	0,14973
4	0,19448
5	0,23687
10	0,41763
15	0,55558
20	0,66085
30	0,80249
40	0,88498
50	0,93301
70	0,97728
100	0,99551

2.3 lentelės tikimybės grafiškai atvaizduotos 2.3 paveiksle:



2.3 pav. Apskaičiuotos įvykio “ypatingai didelis potvynis” pasirodymo tikimybės

Prieduose parodytos pasikartojimo tikimybės, kad per stebima laikotarpį bent vieną kartą bus užtvindyta 45 000 ha teritorija.

2.2.3 PERVIRŠIO TIKIMYBĖS

Perviršio tikimybių metodu apskaičiuojama kiek stebėjimų ateityje viršys praeities stebėjimus.

Visi stebėjimai imtyje surikiuojami mažėjimo tvarka X_1, X_2, \dots, X_n ,

čia X_1 - didžiausia reikšmė, X_n - mažiausiai reikšmė.

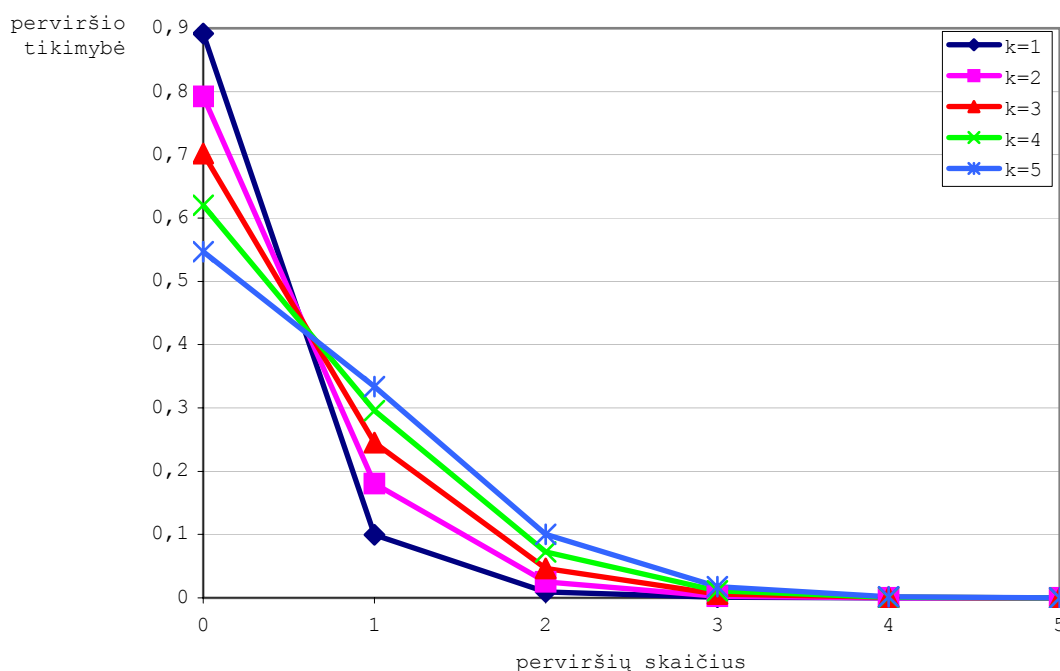
Apskaičiuotos tikimybės per ateinančių 5 metų laikotarpį j kartų viršyti reikšmę X_k yra parodytos 2.4 lentelėje ir 2.4 pav.

2.4 lentelė

5 metų laikotarpio perviršių tikimybės

p	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
$j = 0$	0,89130	0,79227	0,70224	0,62058	0,54670
$j = 1$	0,09903	0,18006	0,24497	0,29552	0,33336
$j = 2$	0,00900	0,02512	0,04666	0,07208	0,10001
$j = 3$	0,00063	0,00239	0,00569	0,01081	0,01795
$j = 4$	0,00003	0,00015	0,00043	0,00097	0,00189
$j = 5$	0,00000	0	0,00002	0,00004	0,00009

Šioje lentelėje $j = 0$ eilutėje parodytos tikimybės, kad k -ojo lygio reikšmė nebus viršijama, $j = 1$ eilutėje – tikimybės, kad nurodyto lygio reikšmė penkerių metų laikotarpyje bus viršijama vieną kartą ir t.t.



2.4 pav. 5 metų laikotarpio perviršių tikimybės

Prieduose pateiktos apskaičiuotos perviršių tikimybės 10 metų laikotarpiui.

2.3 KINTAMOJO *RUSNE* LAIKO EILUČIŲ MODELIO PARINKIMAS

Šioje tyrimo stadijoje analizei bus naudojama 1961 – 1999 metais Rusnės matavimo poste užfiksuotų vandens lygių imtis. 2000 ir 2001 metų stebėjimai į imtį neįtraukiami, nes kintamajam parinkus laiko eilučių modelį, bus patikrinta, ar šie stebėjimai patenka į prognozuojamų reikšmių pasikliautinąjį intervalą.

Tam, kad procesui būtų galima sudaryti laiko eilučių modelį, būtina išsiaiškinti, ar jis nėra grynai atsitiktinis. Tam yra atliekamas hipotezės apie baltąjį triukšmą patikrinimas, kurio rezultatas pateiktas 2.5 lentelėje.

2.5 lentelė

Baltojo triukšmo hipotezės tikrinimas kintamojo *rusne* procesui

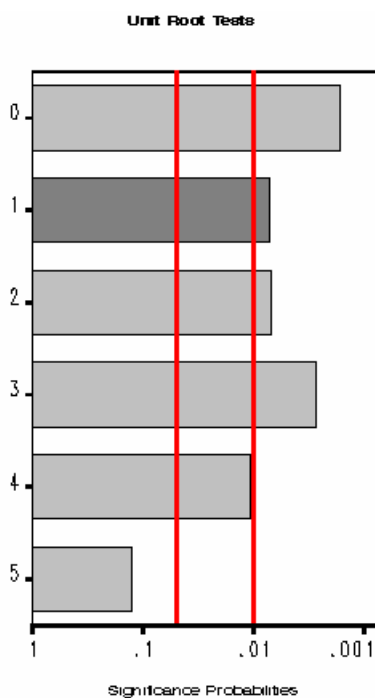
Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	17.33	6	0.0081	0.345	0.064	-0.134	-0.346	-0.324	-0.070

Baltojo triukšmo hipotezės tikrinimas atliekamas proceso autokovariacinėms funkcijoms. Kaip matyti iš 2.5 lentelėje pateikiamų rezultatų, apskaičiuota p -reikšmė yra $0,0081 < 0,05$, todėl hipotezę apie kintamojo *rusne* proceso baltąjį triukšmą galima atmesti.

Kitas žingsnis nagrinėjant laiko eilutes – nustatyti laiko eilutės modelio eilę. Tam yra naudojama statistinio paketo SAS/ETS posistemės procedūra ARIMA bei SAS Time Series Viewer ir SAS Time Series Forecasting System.

Laiko eilučių modelio eilės alternatyvų parinkimas atliekamas dviem metodais: mažiausios kanoninės koreliacijos (SCAN) ir išplėtosios imties autokoreliacinės funkcijos (ESACF) metodais. Modelio eilė nustatoma esant reikšmingumo lygmeniui $\alpha = 0,05$.

Pirmoji nagrinėjama alternatyva - $(p+d,q)=(1,0)$, čia p – autoregresijos eilė, d – integruotumo eilė. Šiam procesui atlikus Dickey – Fuller’io testą, esant autoregresijos eilei $p=1$ (2.5 pav.), buvo nustatyta, kad procesas yra stacionarus ($0,071 < 0,05$). Todėl jam galima pritaikyti AR(1) modelį.



2.5 pav. Dickey – Fuller’io testas $(p+d,q)=(1,0)$

Sudarytam AR(1) modeliui mažiausių kvadratų metodu apskaičiuavus parametrų įverčius, atliktas hipotezės apie liekanų baltojo triukšmo procesą patikrinimas (2.6 lentelė).

2.6 lentelė

Baltojo triukšmo hipotezės tikrinimas AR(1) modelio liekanoms

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	6.84	5	0.2325	0.019	0.016	-0.114	-0.273	-0.243	0.038
12	14.14	11	0.2256	0.029	0.204	0.103	-0.163	0.002	-0.228
18	26.84	17	0.0605	-0.165	-0.070	0.089	0.001	0.369	0.064
24	33.12	23	0.0790	-0.051	-0.040	-0.044	-0.216	0.111	0.036

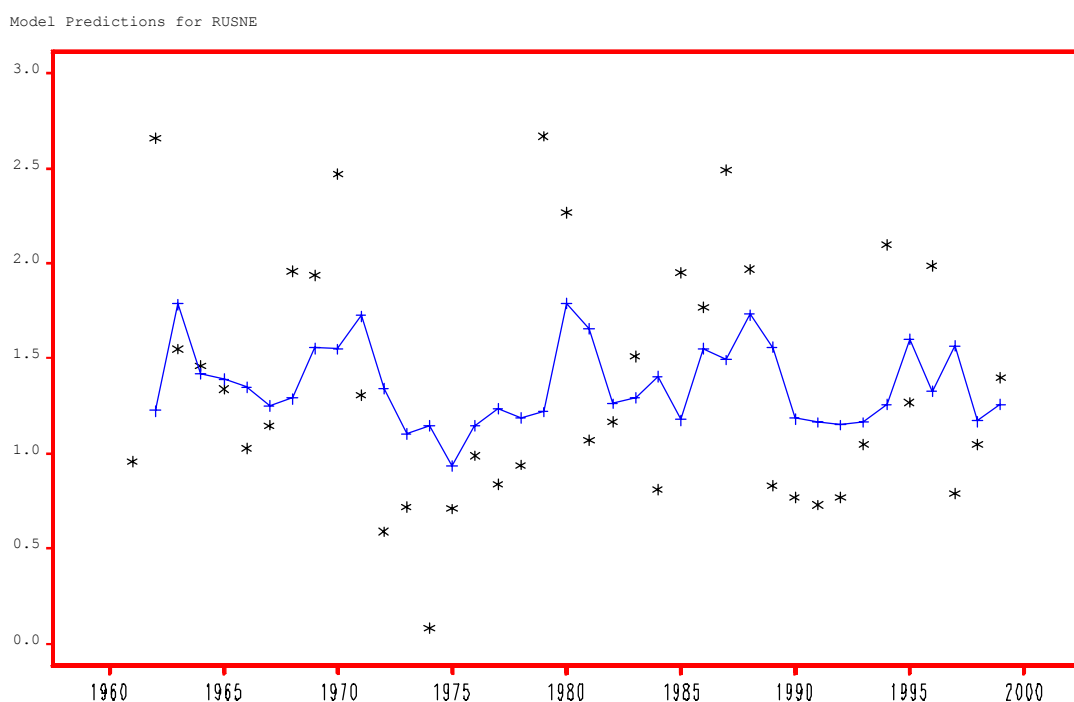
Lentelėje 2.6 pateikti rezultatai rodo, kad sudaryto AR(1) modelio kintamajam *rusne* liekanos yra grynai atsitiktinės, nes apskaičiuotos p -reikšmės visiems suvėlinimams yra didesnės už nustatytą reikšmingumo lygmenį 0,05. Vadinasi sudarytas modelis kintamajam *rusne* tinka.

Modelis AR(1):

$$\xi_t = \xi_0 + a_1 \xi_{t-1},$$

čia $\xi_0 = 1,35783$, $a_1 = 0,32934$.

2.6 paveiksle taškais parodytos stebėtos kintamojo *rusne* reikšmės, o laužtė rodo šių reikšmių aproksimaciją modeliu AR(1).



2.6 pav. AR(1) modelis kintamajam *rusne*

Kitas nagrinėjamas procesas – $(p+d,q)=(0,4)$, t.y. ketvirtos eilės slenkamojo vidurkio modelis MA(4). Kintamojo *rusne* procesas yra stacionarus, todėl jam galima pritaikyti MA(4) modelį.

Mažiausių kvadratų metodu įvertinus modelio parametrus, atliktas baltojo triukšmo hipotezės patikrinimas liekanoms (2.7 lentelė).

2.7 lentelė

Baltojo triukšmo hipotezės tikrinimas MA(4) modelio liekanoms

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	17.45	6	0.0078	0.334	0.059	-0.185	-0.370	-0.318	-0.043
WARNING: The model defined by the new estimates is unstable. The iteration process has been terminated.									
WARNING: Estimates may not have converged.									

Paiškėjo, kad parinkto MA(4) modelio liekanos kintamajam *rusne* nėra atsitiktinės. Be to, modelio įverčiai nekonverguoja, todėl jis nėra tinkamas.

Trečias modelis kintamojo *rusne* stebėjimų procesui – $(p+d,q)=(2,0)$. Esant autoregresijos eilei $p=2$, procesas yra stacionarus, todėl nagrinėjamas modelis yra antros eilės autoregresijos modelis AR(2).

Sudaryto modelio liekanom atliekamas baltojo triukšmo hipotezės tikrinimas (2.8 lentelė). Liekanos yra atsitiktinės, nes apskaičiuotos p -reikšmės yra mažesnės už nustatytą reikšmingumo lygmenį 0,05, todėl galima padaryti išvadą, kad sudarytas AR(2) modelis yra adekvatus.

2.8 lentelė

Baltojo triukšmo hipotezės tikrinimas AR(2) modelio liekanoms

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	6.17	4	0.1871	-0.032	0.070	-0.098	-0.246	-0.237	0.041
12	13.25	10	0.2101	-0.004	0.179	0.095	-0.175	0.014	-0.234
18	25.74	16	0.0578	-0.147	-0.077	0.105	-0.025	0.368	0.049
24	31.99	22	0.0776	-0.032	-0.035	-0.021	-0.218	0.123	0.016

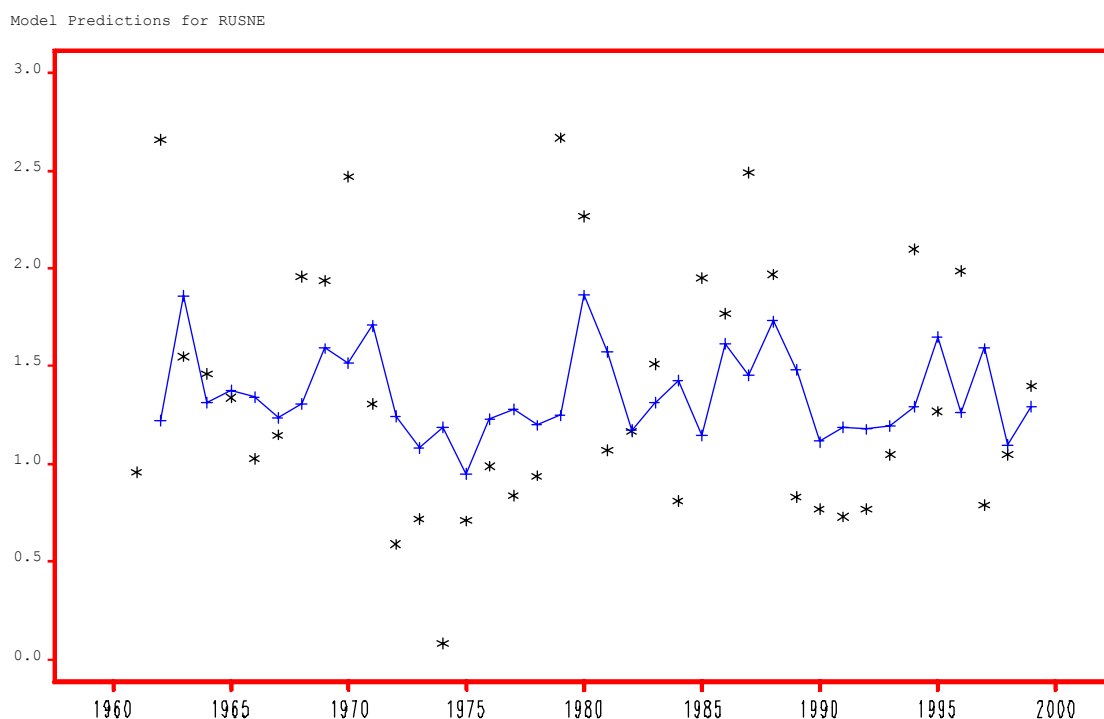
Modelis AR(2):

$$\xi_t = \xi_0 + a_1 \xi_{t-1} + a_2 \xi_{t-2},$$

čia $\xi_0 = 1,3559$,

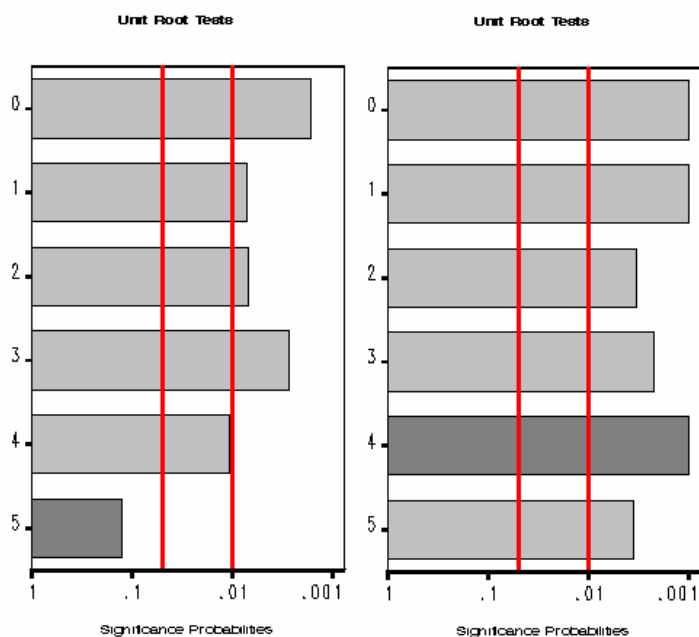
$$a_1 = 0,36109,$$

$$a_2 = -0,08357.$$



2.7 pav. AR(2) modelis kintamajam *rusne*

Paskutinis kintamajam *rusne* nagrinėjamas procesas yra $(p+d,q)=(5,1)$. Autoregresijos eilei p ir integruotumo eilei d nustatyti atliekamas Dickey – Fuller’io testas. Esant autoregresijos eilei $p=5$, šis procesas nėra stacionarus (2.8(a) pav.), todėl būtina jį diferencijuoti – apskaičiuoti pirmos eilės pokyčius ir vėl atlikti Dickey – Fuller’io testą (2.8(b) pav.).



a) $p=5, d=0$

b) $p=4, d=1$

2.8 pav. Dickey – Fuller’io testas $(p+d,q)=(5,1)$ procesui

Esant autoregresijos eilei $p=4$ ir integruotumo eilei $q=1$, procesas yra stacionarus. Vadinasi jis yra 1-os eilės integruotas, nes pirmos eilės pokyčiai yra stacionarūs ($0,0004 < 0,05$). Taigi, sudarytas modelis yra ARIMA(4,1,1).

Modelio parametrai yra įvertinami mažiausių kvadratų metodu. Liekanoms atliekamas baltojo triukšmo hipotezės tikrinimas, kurio rezultatai pateikti 2.9 lentelėje:

2.9 lentelė

Baltojo triukšmo hipotezės tikrinimas ARIMA(4,1,1) modelio liekanoms

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	6.59	6	0.3606	-0.261	-0.045	-0.048	-0.181	-0.177	0.122
WARNING: The model defined by the new estimates is unstable. The iteration process has been terminated.									
WARNING: Estimates may not have converged.									

Kaip matyti iš 2.9 lentelėje pateiktų rezultatų, sudaryto ARIMA(4,1,1) modelio liekanos nėra grynai atsitiktinės, o sudaryti parametrai įverčiai diverguoja. Todėl šis modelis nėra tinkamas kintamojo *rusne* reikšmėms prognozuoti.

Taigi, išanalizavus visus pasiūlytus variantus kintamojo *rusne* procesui, buvo atrinkti 2 modeliai: AR(1) ir AR(2). Tam, kad išsiaiškintume, kuris modelis yra tikslesnis, reikia palyginti modelių prognozės tikslumo matą – vidutinę kvadratinę paklaidą (2.10 lentelė).

2.10 lentelė

Kintamojo *rusne* sudarytų alternatyvių modelių paklaidų palyginimas

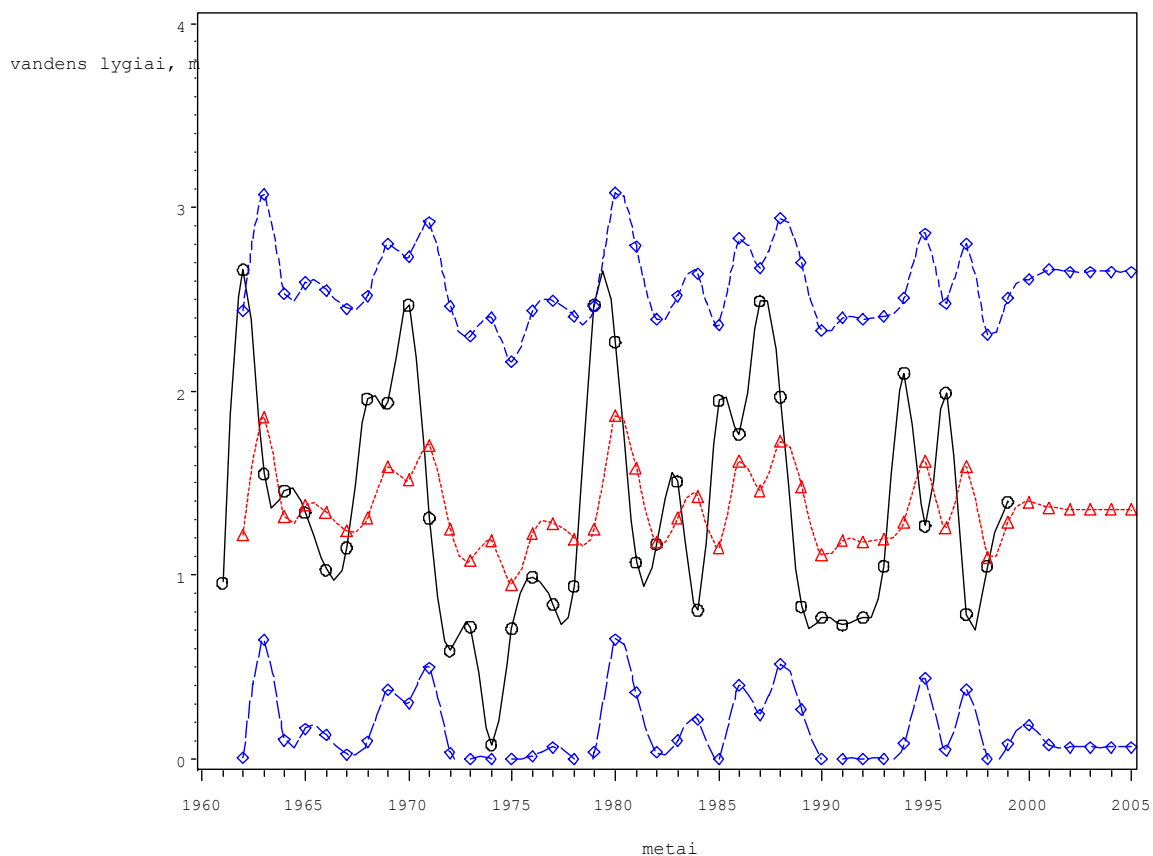
Paklaidos	MODELIS			
	AR(1)		AR(2)	
	Su laisvuju nariu	Be laisvojo nario	Su laisvuju nariu	Be laisvojo nario
Vidutinė kvadratinė paklaida	0,36112	0,51424	0,35917	0,49274
Vidutinė absoliutinė paklaida	0,48421	0,55125	0,48031	0,53099

Kintamajam *rusne* modelis AR(2) yra tikslesnis, nes vidutinė kvadratinė ir vidutinė absoliutinė paklaidos jam yra mažesnės.

Įvertinus visus alternatyvius modelius, kintamajam *rusne* parenkamas AR(2) modelis:

$$\xi_t = 1,3559 + 0,36109 \cdot \xi_{t-1} - 0,08357 \cdot \xi_{t-2}.$$

2.9 paveiksle juodąja kreive parodytos tikrosios kintamojo *rusne* stebėjimų reikšmės, raudonąja kreive – modelio AR(2) apskaičiuotos reikšmės, o mėlynomis kreivėmis – modelio apskaičiuotų reikšmių pasikliautinųjų intervalų rėžiai, esant pasiklivimo lygmeniui 0,95.



2.9 pav. AR(2) modelis kintamajam *rusne*

Patikrinkime, ar kintamajam *rusne* sudarytas modelio prognozė atitinka 2000 ir 2001 metais stebėtas reikšmes.

2.11 lentelė

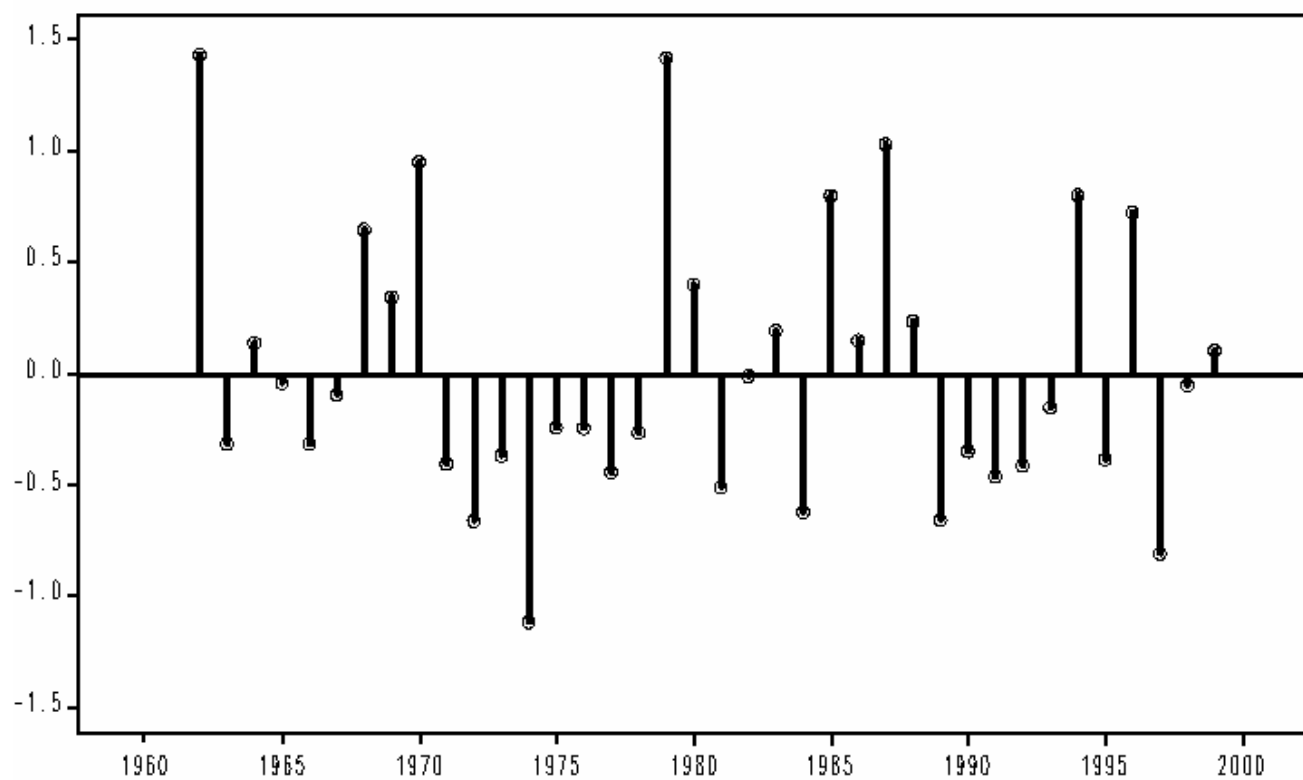
Modelio AR(2) sudarytų prognozių patikrinimas

Metai	Stebėta reikšmė	Pasikliautinis intervalas, esant pasiklovimo lygmeniui 0,95
2000	0,62	(0,1851; 2,6097)
2001	0,53	(0,0783; 2,6561)

Iš 2.11 lentelės matyti, kad abi reikšmės patenka į sudarytą pasikliautinąjį intervalą, todėl galima teigti, kad parinktas modelis pakankamai tiksliai prognozuoja Nemuno vandens lygį matavimo poste Rusnė.

2.10 paveiksle pateiktos kintamojo *rusne* prognozavimo paklaidos, apskaičiuotos pritaikius AR(2) modelį.

Prediction errors for RUSNE

**2.10 pav. Prognozavimo paklaidos, apskaičiuotos kintamajam *rusne*, pritaikius AR(2) modelį**

3. PROGRAMINĖ REALIZACIJA IR INSTRUKCIJA VARTOTOJUI

Darbas su programa POTVYNIAI v 1.0

Programa sudaryta C++ Builder v 5.0 pagrindu ir skirta apskaičiuoti perviršio bei pasikartojimo tikimybes, momentinius stebėjimo lygių rekordus. Galima peržiūrėti tekstinius duomenų failus bei papildyti juos naujais stebėjimais.

Viršutinis meniu:

1. Failas:

- Baigti darbą – darbo su programa pabaigos pasirinkimas;

2. Pagalba:

- Aprašymas – darbo su programa aprašymas (Help);
- Apie programą – programos paskirtis.

1. Kortelė "Lygiai"

Kiekvienam iš trijų vandens lygių matavimo postų (Smalininkai, Tilžė, Rusnė) apskaičiuojamas kiekvieno stebėjimo momentinis rekordas, kuris saugomas kartu su vandens lygių stebėjimais duomenų faile *data.txt*.

Duomenų įvedimas:

1) Metai: įvedama reikšmė turi būti sveikas skaičius, didesnis už 1961. Klaidingai įvedus duomenis pasirodo įspėjamasis pranešimas.

2) Vandens lygis, m: įvedamas nurodytais metais stebėtas didžiausias vandens lygis atitinkamame matavimo poste; įvedamos reikšmės sveikoji dalis turi būti atskiriama tašku.

Įvedę metus, Tabuliacijos klavišo arba pelės pagalba pereikite prie vandens lygio įvedimo laukų. Užpildę visus laukus paspauskite mygtuką "Saugoti" – įvesti duomenys bus įrašyti į potvynių lygių duomenų failą *data.txt*, kurį galite peržiūrėti paspaudę mygtuką "Rodyti".

Kiekvienam stebėjimo postui yra užrašyti stebėti vandens lygiai ir jų momentiniai rekordai.

Perviršių tikimybės

Apskaičiuojamos perviršių tikimybės.

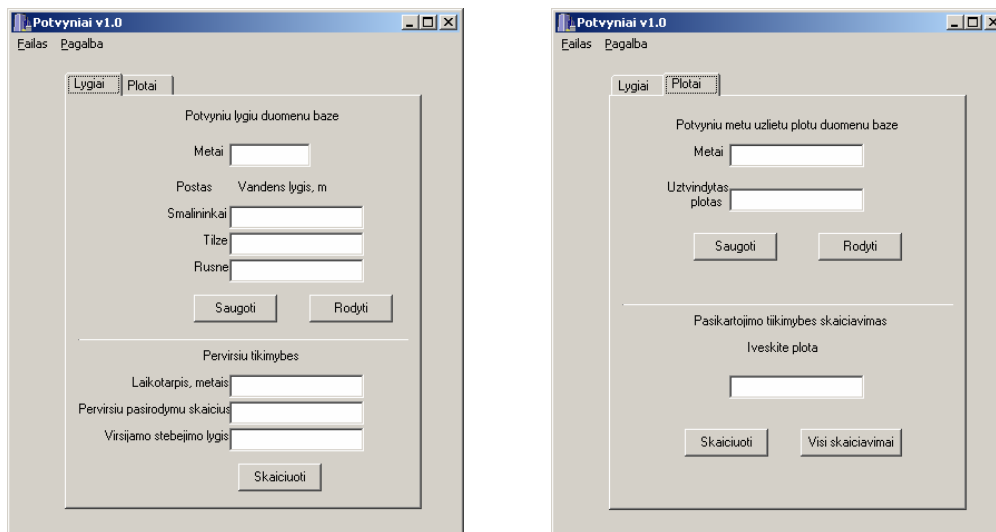
Duomenų įvedimas:

1) Laikotarpis, metais: įvedamas sveikas skaičius – kelerių metų laikotarpiui apskaičiuojamos perviršio tikimybės.

2) Perviršių pasirodymų skaičius: įvedamas sveikas skaičius, rodantis kiek kartų per pasirinktą laikotarpį viršijamas stebėjimo lygis. Ši reikšmė turi būti mažesnė už pasirinktą laikotarpį.

3) Viršijamo stebėjimo lygis: įvedamas sveikas skaičius, kuris parodo kurio lygio stebėjimui skaičiuojama perviršio tikimybė. Ši reikšmė turi būti mažesnė už duomenų faile saugomų stebėjimų skaičių.

Įvedę duomenis, paspauskite mygtuką "Skaičiuoti" – bus apskaičiuota su kokia tikimybe pasirinktas stebėjimo lygis per tam tikrą laikotarpį viršijamas nurodytą skaičių kartų.



3.1 pav. Programos langas

2. Kortelė "Plotai"

Apskaičiuojama tikimybė, kad per tam tikrą laikotarpį bus užtvindyta vienodo dydžio teritorija.

Duomenų įvedimas:

1) **Metai:** įvedama reikšmė turi būti sveikas skaičius, didesnis už 1961.

2) **Užtvindytas plotas:** įvedamas nurodytais metais užtvindytos teritorijos plotas. Skaičiaus sveikoji dalis turi būti atskiriama tašku;

Įvedę duomenis, paspauskite mygtuką "Saugoti" - įvesti duomenys bus užsaugoti potvynių metu užlietų plotų duomenų faile *plotas.txt*, kuri galima peržiūrėti paspaudus mygtuką "Rodyti".

Pasikartojimo tikimybės skaičiavimas

Apskaičiuojamos tikimybės, kad potvynio metu pakartotinai bus užtvindytas įvestas plotas.

Duomenų įvedimas:

'**Įveskite plotą**' – įvedama užlieto ploto reikšmė, kuriam norima apskaičiuoti pasikartojimo tikimybę. Jeigu įvesto užtvindyto ploto duomenų faile nerasta, tai atsiranda pranešimas "Nėra vienodų įvykių".

Paspaudus mygtuką "Skaičiuoti" atidaromas duomenų failas *rk.txt*, kuriame parodoma kuriais metais buvo užlietas vienodas nurodytas plotas ir apskaičiuotos tikimybės, kad toks plotas bus užlietas per 1-100 metų laikotarpį bei nubraižomas pasikartojimo tikimybių grafikas.

IŠVADOS

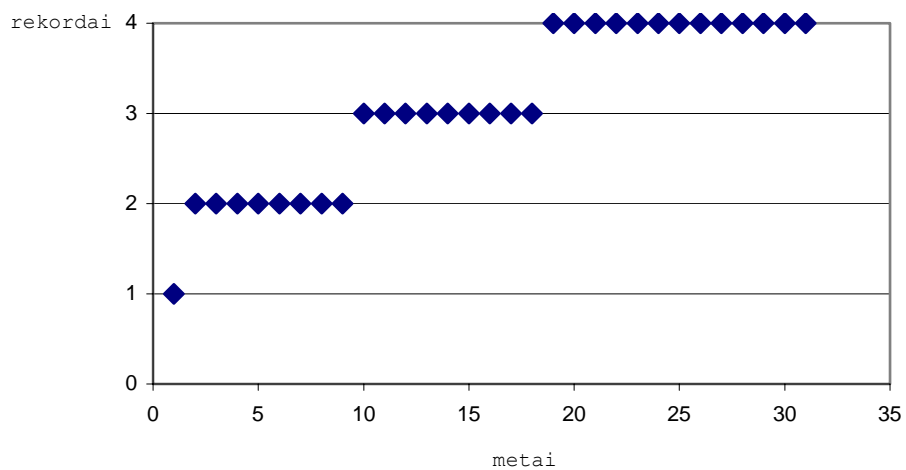
Susipažinus su užsienyje taikomais ekstremalių įvykių tyrimo metodais ir pritaikius juos potvynių Nemuno deltoje tyrimui, gauti tokie rezultatai:

1. Naudojant rekordų metodą, nustatyta, kad tirti kintamieji – *smalininkai*, *tilze*, *rusne* – yra nepriklausomi ir vienodai pasiskirstę.
2. Taikant perviršių metodą buvo apskaičiuotos perviršių tikimybės penkerių ir dešimties metų laikotarpiams.
3. Taikant pasikartojimo periodo metodą, apskaičiuotos tam nusakyto įvykio “ypatingai didelis potvynis” pasikartojimo tikimybės šimto metų laikotarpiui.
4. Panaudojus C++ Builder 5.0 sukurta programa, kuri leidžia vartotojui apskaičiuoti aukščiau minėtas tikimybes ir saugoti gautus rezultatus. Taip galima papildyti turimas duomenų bazes naujais stebėjimais ir perskaičiuoti tikimybes naujiems duomenims.
5. Kiekvienam kintamajam buvo sudaryti keli laiko eilučių modeliai, iš kurių atrinktas geriausiai atitinkantis stebėjimus. Kintamajam *rusne*: AR(2), kintamajam *smalininkai*: AR(1).

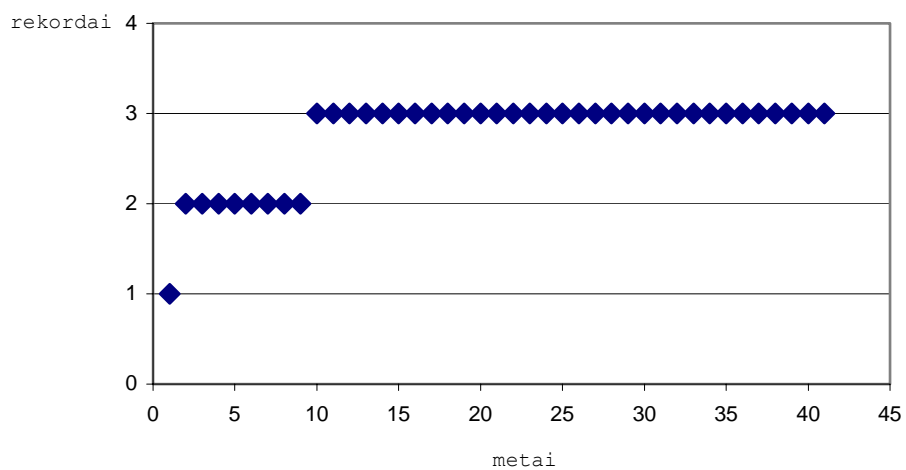
LITERATŪROS SĄRAŠAS

1. V. Čekanavičius, G. Murauskas. Statistika ir jos taikymai, I. Vilnius: TEV, 2001. – 239 p.
2. V. Čekanavičius, G. Murauskas. Statistika ir jos taikymai, II. Vilnius: TEV, 2002. – 271 p.
3. N. Kligienė. Įvadas į atsitiktinių sekų statistinę analizę, Vilnius, Technika, 1998. – 138 psl.
4. Ataskaita: Pavasario potvynių dinamika Nemuno deltoje, Vilnius, TSRS Vyriausioji hidrometeorologinės tarnybos valdyba, 1973. – 123 psl.
5. Embrechts P., Kluppelberg C., Mikosch T. Modelling Extremals Events, Springer-Verlag Berlin Heidelberg 1997. – 645 p.
6. B. Gailiušis, J. Jablonskis, M. Kovalenkoviėnė. Lietuvos upės. Hidrografija ir nuotėkis, Kaunas, Lietuvos energetikos institutas, 2001. – 792 psl.
7. A. Vidžiūnas. C++ ir C++ Builder pradmenys, Kaunas, Smaltija, 2002. – 316 psl.
8. J. Blonskis, V. Bukšnaitis, V. Jusas, R. Marcinkevičius, J. Smolinskas. Programavimo C++Builder pavyzdžiai, Kaunas, Smaltija, 2003. – 292 psl.
9. Rebecca J. Elliott – Learning SAS in the Computer Lab, Duxbury Press, 1995. – 175p.
10. Статистические методы в гидрологии – Гидрометеорологическое издательство, Ленинград, 1970. – 270стр.

1 PRIEDAS. REKORDINIAI MOMENTAI



1.1 pav. Kintamojo *tilze* rekordiniai momentai



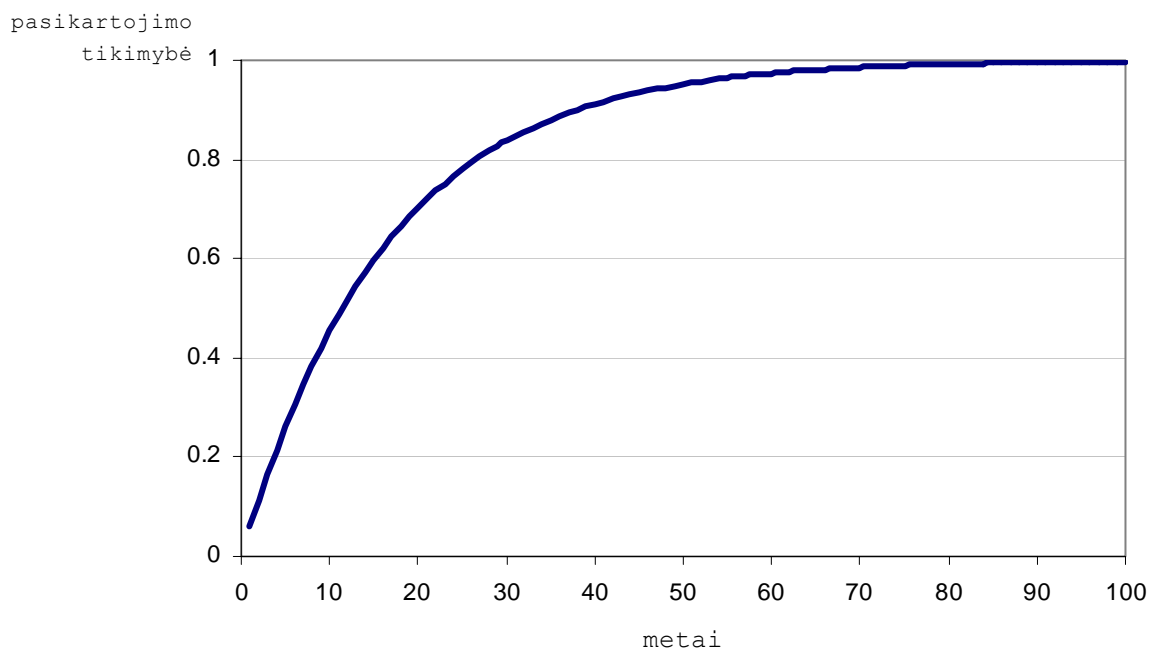
1.2 pav. Kintamojo *smalininkai* rekordiniai momentai

2 PRIEDAS. PASIKARTOJIMO TIKIMYBĖS

2.1 lentelė

Tikimybė bent vieną kartą k metų laikotarpyje potvynio metu užtvindyti 45 000 ha teritoriją

k	r_k
1	0,05882
2	0,11419
3	0,16629
4	0,21534
5	0,26149
10	0,45461
15	0,59722
20	0,70255
30	0,83777
40	0,91152
50	0,95174
70	0,98565
100	0,99767



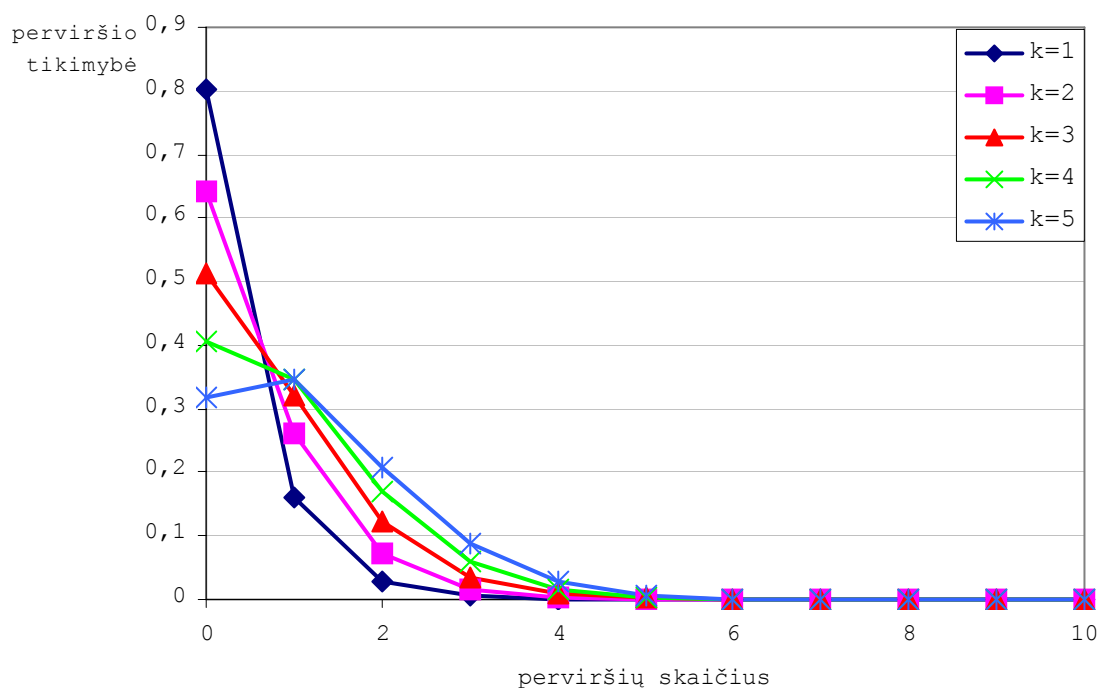
2.1 pav. Tikimybė, kad bent vieną kartą per stebimą laikotarpį potvynio metu bus užtvindyta 45 000 ha teritorija

3 PRIEDAS. PERVIRŠIŲ TIKIMYBĖS

3.1 lentelė

10 metų laikotarpio perviršių tikimybės

p	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
$j=0$	0,80392	0,64314	0,51188	0,40524	0,31902
$j=1$	0,16078	0,26251	0,31993	0,34489	0,34676
$j=2$	0,02953	0,07383	0,12253	0,16869	0,20806
$j=3$	0,00492	0,01676	0,03551	0,05998	0,08827
$j=4$	0,00073	0,00319	0,00829	0,01670	0,02874
$j=5$	0,00010	0,00051	0,00158	0,00373	0,00739
$j=6$	0,00001	0,00007	0,00025	0,00067	0,00150
$j=7$	0	0,00001	0,00003	0,00009	0,00024
$j=8$	0	0	0	0,00001	0,00003
$j=9$	0	0	0	0	0
$j=10$	0	0	0	0	0

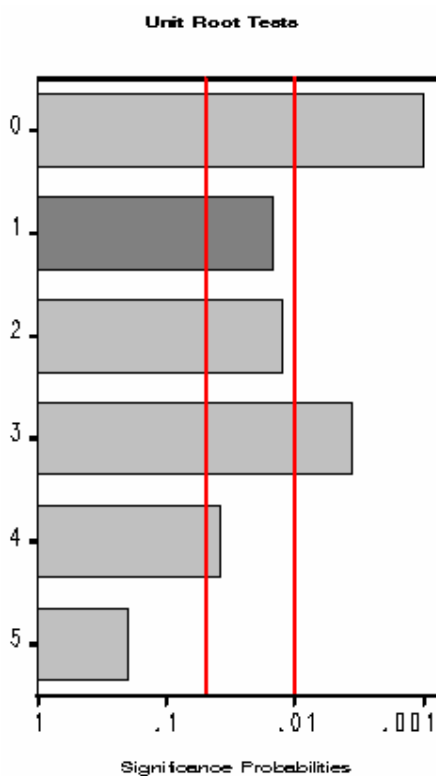


3.1 pav. 10 metų laikotarpio perviršių tikimybės

4 PRIEDAS. LAIKO EILUČIŲ MODELIO KINTAMAJAM *SMALININKAI* PARINKIMAS

Modelio eilė kintamojo *smalininkai* procesui nustatoma dviem būdais: mažiausios kanoninės koreliacijos (SCAN) ir išplėtosios imties autokoreliacinės funkcijos (ESACF) metodais. Toliau nagrinėjami ESACF bei SCAN metodų pasiūlyti alternatyvūs modeliai.

Pirmasis jų $(p+d, q)=(1,0)$ procesas, čia p – autoregresijos eilė, d – integruotumo eilė. Atlikus Dickey – Fuller’io testą (4.1 pav.) gauta, kad procesas AR(1) yra stacionarus, nes apskaičiuota tikimybė 0,01477 yra mažesnė už nustatytą reikšmingumo lygmenį 0,05.



$$p=1, d=0$$

4.1 pav. Dickey – Fuller’io testas $(p+d, q)=(1,0)$ procesui

Mažiausių kvadratų metodu įvertinus modelio AR(1) parametrus, atliktas baltojo triukšmo testas modelio paklaidoms (4.1 lentelė).

4.1 lentelė

Baltojo triukšmo hipotezės tikrinimas AR(1) modelio liekanoms

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	5.04	5	0.4108	-0.010	0.076	-0.092	-0.281	-0.106	0.076
12	15.43	11	0.1638	-0.039	0.255	0.063	-0.279	-0.073	-0.191
18	25.36	17	0.0870	-0.060	-0.142	0.178	0.063	0.286	0.034
24	33.56	23	0.0717	-0.037	-0.023	0.014	-0.201	0.176	0.104

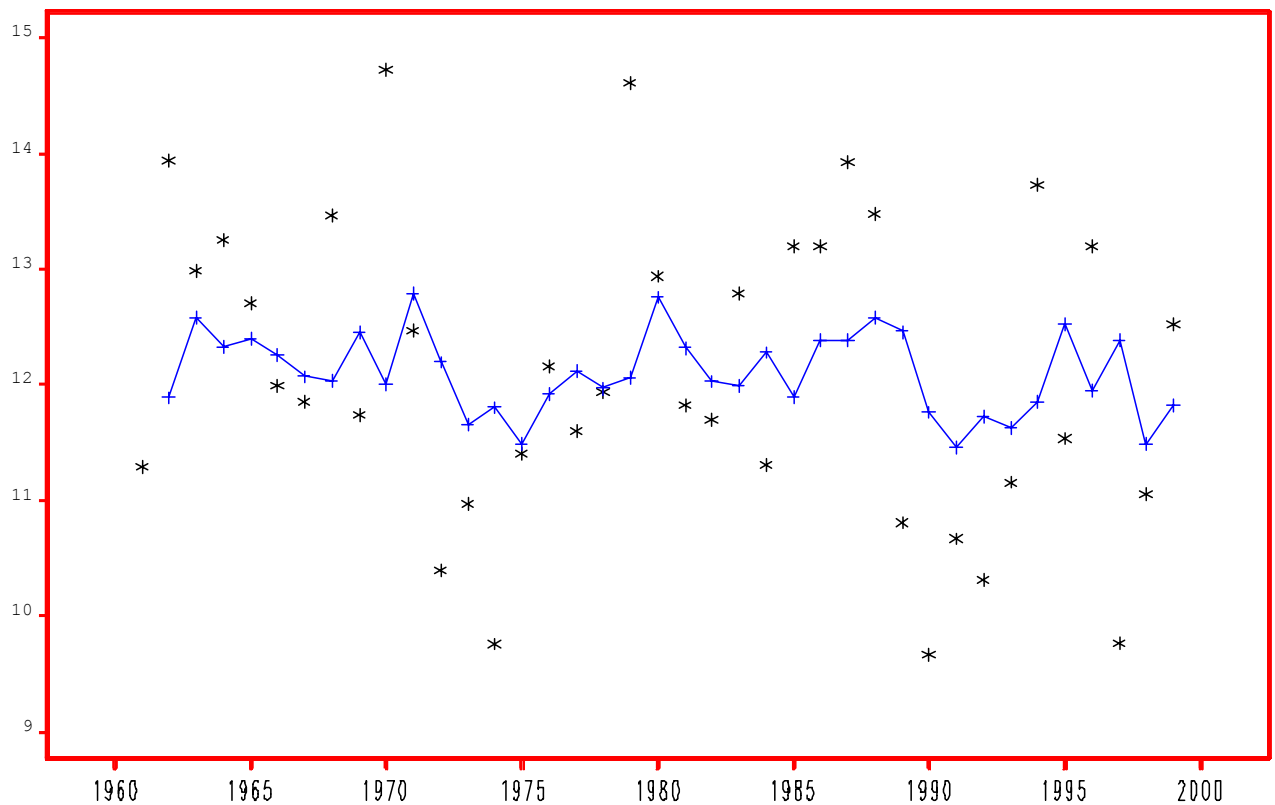
Apskaičiuotos χ^2 testo statistikos rodo, kad liekanos, apskaičiuotos modeliui AR(1), yra atsitiktinis procesas, nes apskaičiuotos p -reikšmės yra didesnės už nustatytą reikšmingumo lygmenį 0,05. Taigi, galima padaryti išvadą, kad sudarytas AR(1) modelis (4.2 pav.)

$$\xi_t = \xi_0 + a_1 \xi_{t-1},$$

čia $\xi_0 = 12,1017$, $a_1 = 0,26165$

yra tinkamas.

Model Predictions for SMALININKAI



4.2 pav. AR(1) modelis kintamajam *smalininkai*

Kitas alternatyvus pasiūlytas proceso modelis – $(p+d, q)=(0,4)$, t.y. 4-os eilės slenkamojo vidurkio modelis MA(4). Šis procesas yra stacionarus, todėl jam galima taikyti laiko eilutės modelį.

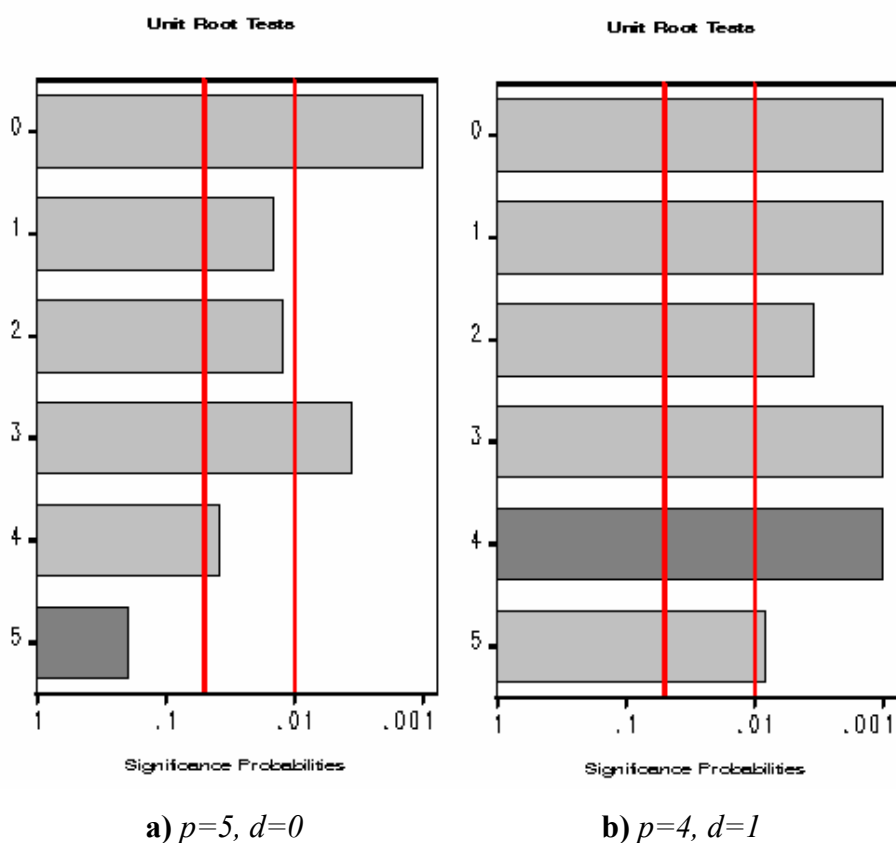
Modelio parametrai įvertinami mažiausių kvadratų metodu. Ir sudaryto modelio liekanoms atliekamas hipotezės apie baltąjį triukšmą patikrinimas (4.2 lentelė).

Baltojo triukšmo hipotezės tikrinimas MA(4) modelio liekanoms

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	10.08	6	0.1213	0.265	0.100	-0.131	-0.319	-0.162	0.031
WARNING: The model defined by the new estimates is unstable. The iteration process has been terminated.									
WARNING: Estimates may not have converged.									

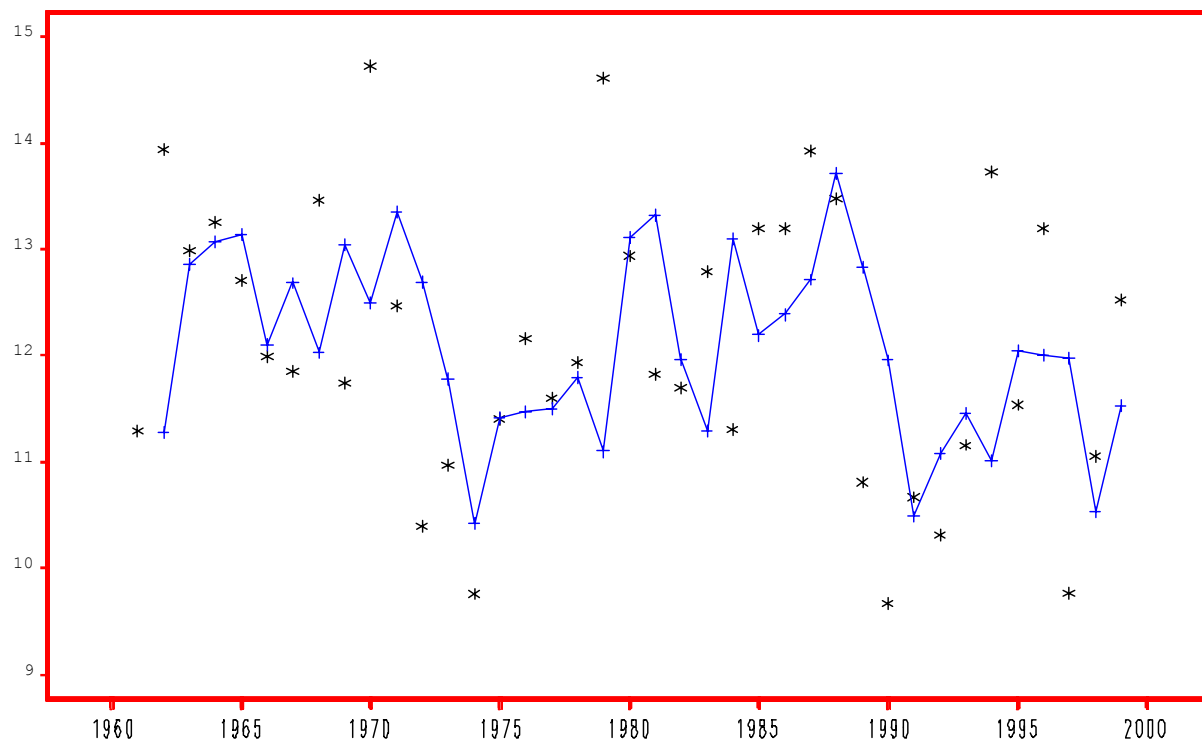
Kaip matyti iš 4.2 lentelės, sudaryto modelio įverčiai nekonverguoja, todėl jis šiam procesui aprašyti nėra tinkamas.

Trečias alternatyvus procesas kintamajam *tilze* - $(p+d, q)=(5,0)$. Autoregresijos eilei p ir integruotumo eilei d nustatyti naudojamas Dickey – Fuller’io testas, kurio rezultatai parodyti 4.3 pav. Esant autoregresijos eilei $p=5$, procesas nėra stacionarus ($0,1953 > 0,05$). Tačiau pirmos eilės pokyčiai yra stacionarūs, vadinasi procesas yra 1-os eilės integruotas, esant autoregresijos eilei $p=4$ apskaičiuota tikimybė lygi $0,0004 < 0,05$.



4.3 pav. Dickey – Fuller’io testas $(p+d,q)=(5+0,0)$ procesui

Model Predictions for SMALININKAI



4.4 pav. ARIMA(4,1,0) modelis kintamajam *smalininkai*

Modelio ARIMA(4,1,0) parametrai buvo įvertinti mažiausių kvadratų metodu, o apskaičiuotoms liekanoms atliktas baltojo triukšmo testas (4.3) lentelė

4.3 lentelė

Baltojo triukšmo hipotezės tikrinimas ARIMA(4,1,0) modelio liekanoms

The ARIMA Procedure									
Autocorrelation Check of Residuals									
To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	2.99	2	0.2240	-0.104	-0.036	-0.112	-0.068	-0.194	-0.007
12	14.78	8	0.0636	-0.060	0.240	0.176	-0.348	-0.062	-0.068
18	27.86	14	0.0149	-0.024	-0.217	0.194	0.044	0.317	-0.046
24	32.65	20	0.0368	-0.018	-0.008	0.029	-0.185	0.059	0.100

Iš 4.3 lentelėje pateiktų baltojo triukšmo testo rezultatų matyti, kad pirmiems 12 vėlinimams liekanos yra atsitiktinis procesas.

Tačiau, palyginus 4.1 ir 4.3 lentelių rezultatus galima pastebėti, kad AR(1) modelio paklaidų baltojo triukšmo testas tenkinamas didesne tikimybe. Be to, palyginus sudarytų modelių prognozės tikslumo matą – vidutinę kvadratinę paklaidą – bei kitas paklaidas (4.4 lentelė) matyti, kad AR(1) geriau atitinka stebėtas reikšmes.

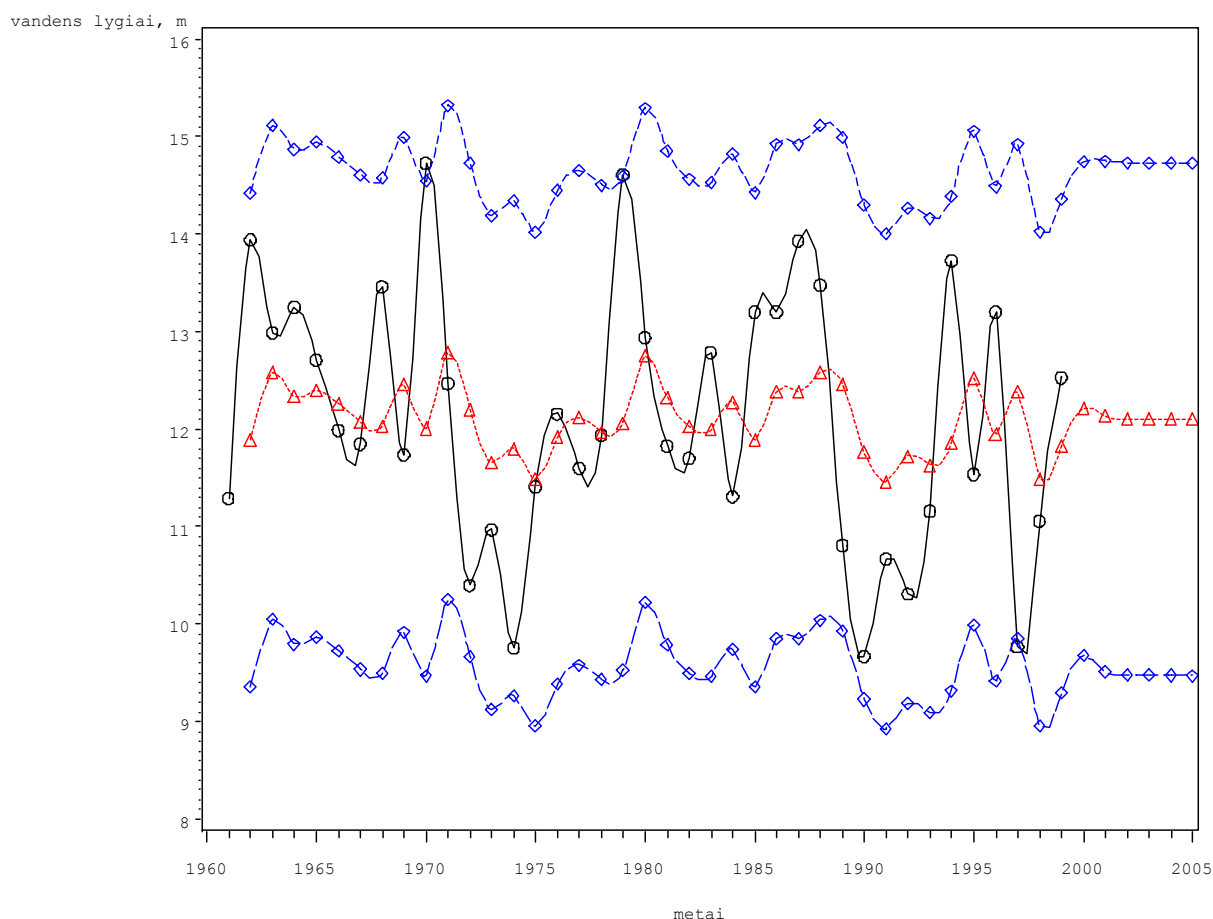
Kintamojo *smalininkai* sudarytų alternatyvių modelių paklaidų palyginimas

Paklaidos	Modelis			
	AR(1)		ARIMA(4,1,0)	
	Su laisvuju nariu	Be laisvojo nario	Su laisvuju nariu	Be laisvojo nario
Vidutinė kvadratinė paklaida	1,61243	2,55085	1,91962	1,91996
Vidutinė absoliutinė paklaida	1,02425	1,33140	1,06887	1,07130

Taigi, įvertinus visus alternatyviuosius modelius, kintamajam *smalininkai* parenkamas AR(1) modelis:

$$\xi_t = 12,1017 + 0,26165 \cdot \xi_{t-1}$$

4.5 paveikslėlyje juodąja kreive parodytos tikrosios kintamojo *tilze* stebėjimų reikšmės, raudonąja kreive – modelio AR(1) apskaičiuotos reikšmės, o mėlynomis kreivėmis – modelio apskaičiuotų reikšmių pasikliautinųjų intervalų rėžiai, esant pasikliovimo lygmeniui 0,95.



4.5 pav. AR(1) modelis kintamajam *smalininkai*

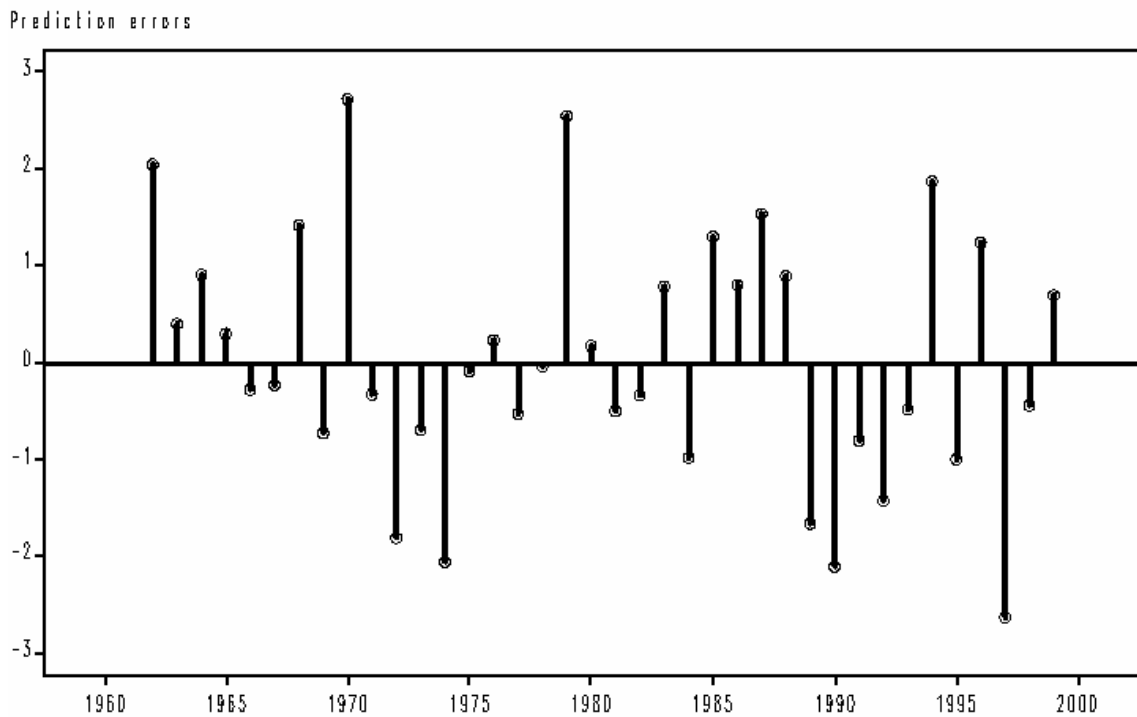
Patikrinkime, ar kintamajam *smalininkai* sudarytas modelio prognozė atitinka 2000 ir 2001 metais stebėtas reikšmes.

Modelio AR(1) sudarytų prognozių patikrinimas

Metai	Stebėta reikšmė	Pasikliautinis intervalas, esant pasiklydimo lygmeniui 0,95
2000	10,01	(9,6789; 14,7485)
2001	9,73	(9,5108; 14,7511)

Iš 4.5 lentelės matyti, kad abi reikšmės patenka į sudarytą pasikliautinąjį intervalą, todėl galima teigti, kad parinktas AR(1) modelis kintamajam *smalininkai* pakankamai tiksliai prognozuoja Nemuno vandens lygį matavimo poste Smalininkai.

4.6 paveiksle parodytos kintamajam *smalininkai* apskaičiuotos prognozavimo paklaidos, pritaikius AR(1) modelį.



4.6 pav. Prognozavimo paklaidos, apskaičiuotos kintamajam *smalininkai* pritaikius AR(1) modelį

5 PRIEDAS. PROGRAMŲ TEKSTAI

5.1 TIKIMYBIŲ SKAIČIAVIMAS

Pagrindinė programa:

```
//-----
#include <vcl.h>
#pragma hdrstop

#include "Potviniai.h"
#include "potviniai.h"
#include "fstream.h"
#include "iostream.h"
#include "string.h"
#include "stdlib.h"
#include "stdio.h"
#include "Isvedimas.h"
#include "pagalba.h"
#include "apie.h"
//-----
#pragma package(smart_init)
#pragma resource "*.dfm"
TForm1 *Form1;
float m[4];
//-----
__fastcall TForm1::TForm1(TComponent* Owner)
    : TForm(Owner)
{
    Edit1->Clear();
    Edit2->Clear();
    Edit3->Clear();
    Edit4->Clear();
    Edit5->Clear();
    Edit6->Clear();
    Edit7->Clear();
    Edit8->Clear();
    Edit9->Clear();
    Edit10->Clear();
}
//*****
// Valdymo mygtuku aptarnavimas
//*****
//-----
// potvyniu lygiu duomeniu bazes mygtukas "Saugoti"
//-----
void __fastcall TForm1::Button1Click(TObject *Sender)
{
    char *failas="data.dat";
    AnsiString zodis[4];
    int sk,poz=0,i=0,nr;
    float lygis[4];
    char eil[30],*rod;

    nr=skaitymas1();
    poz=0; //duomeniu teisingumo pozymis
    ofstream outf (failas, ios:: app); //rasymo srautas
    if (!outf){Application->MessageBox("Nera failo!", "Pranesimas", MB_OK+MB_ICONERROR);}
    zodis[0]=Edit1->Text;
    zodis[1]=Edit2->Text;
    zodis[2]=Edit3->Text;
    zodis[3]=Edit4->Text;
    strcpy(eil, zodis[0].c_str());
    //-----
    //ar duomenys ivesti teisingai?
    //-----
    lygis[0]=strtod(eil, &rod);
```

```

if (atoi(eil)< 1961 || *rod!=NULL){
    Application->MessageBox("Klaidingai ivesti
metai!", "Pranesimas", MB_OK+MB_ICONERROR);
    poz=1;
}
else {
//zodis[0]=IntToStr(atoi(zodis[0]))
for (i=1;i<4;i++){
strcpy(eil,zodis[i].c_str());
lygis[i]=strtod(eil,&rod);
    if (lygis[i]!=0)
        {if (*rod!=NULL) {
            Application->MessageBox("Klaidingai ivestas
lygis!", "Pranesimas", MB_OK+MB_ICONERROR);
            poz=1;
            break;
        }
        else {
            zodis[i]=FloatToStrF(lygis[i],ffFixed,4,3);
        }
    }
else {if (*rod==NULL) zodis[i]=FloatToStrF(lygis[i],ffFixed,4,3);
else {
    Application->MessageBox("Klaidingai ivetsas
lygis!", "Pranesimas", MB_OK+MB_ICONERROR);
    poz=1;
    break;
    }
}
}
}
if (poz==0){
//-----
//ivestu duomenu R skaiciavimas
maksimumas(nr);
if (nr!=2){
    if (lygis[1]>m[0]) B[nr].smal_R++;
    if (lygis[2]>m[1]) B[nr].tilze_R++;
    if (lygis[3]>m[2]) B[nr].rusne_R++;
}
//-----
// duomenu irasymas i faila

for (i=0;i<4;i++)
strcpy(A[i].duomenys,zodis[i].c_str());

outf << A[0].duomenys << "\t";
outf << A[1].duomenys << "\t";
outf << B[nr].smal_R << "\t";
outf << A[2].duomenys << "\t";
outf << B[nr].tilze_R << "\t";
outf << A[3].duomenys << "\t";
outf << B[nr].rusne_R << endl;
Edit1->Clear();
Edit2->Clear();
Edit3->Clear();
Edit4->Clear();

}
outf.close();
}
//-----
// Potvyniu lygiu duomenu bazes mygtukas "Rodyti"
//-----
void __fastcall TForm1::Button2Click(TObject *Sender)
{
char *failas="data.dat"; //duomenu failo pavadinimas
Form2->Memo1->Lines->LoadFromFile(failas);
Form2->Caption="Potvyniu lygiu duomenu baze";
Form2->ShowModal();
}

```

```

//-----
// Potvyniu metu uzlietu plotu duomenu bazes mygtukas "Saugoti"
//-----
void __fastcall TForm1::Button3Click(TObject *Sender)
{
char *failas="plotas.dat";
AnsiString zodis[4];
int sk,poz=0,i=0;
float lygis[4];
char eil[30],*rod;
char d1[30],d2[30],d3[30],d4[30];

    poz=0; //duomenu teisingumo pozymis
    ofstream outf (failas, ios:: app); //rasymo srautas
    if (!outf){Application->MessageBox("Neatidarytas
failas!", "Pranesimas",MB_OK+MB_ICONERROR);}
    zodis[0]=Edit5->Text;
    zodis[1]=Edit6->Text;
    strcpy(eil,zodis[0].c_str());
    //-----
    //ar duomenys ivesti teisingai?
    //-----
    lygis[0]=strtod(eil,&rod);
    if (atoi(eil)< 1961 || *rod!=NULL){
        Application->MessageBox("klaidingi metai!", "Pranesimas",MB_OK+MB_ICONERROR);
        poz=1;
    }
    else {
        strcpy(eil,zodis[1].c_str());
        lygis[1]=strtod(eil,&rod);
        if (lygis[1]!=0)
            {if (*rod!=NULL) {
                Application->MessageBox("Klaidingas
plotas!", "Pranesimas",MB_OK+MB_ICONERROR);
                poz=1;
            }
            else {
                zodis[1]=FloatToStrF(lygis[1],ffFixed,8,3);
                // Memo1->Lines->Append(zodis[i]);
            }
        }
        else {if (*rod==NULL) zodis[1]=FloatToStrF(lygis[1],ffFixed,8,3);
            else {
                Application->MessageBox("Klaidingas plotas!", "Pranesimas",MB_OK+MB_ICONERROR);
                poz=1;
            }
        }
    }
    //-----
    // duomenu irasymas i faila
    if (poz==0){
        for (i=0;i<2;i++)
            strcpy(A[i].duomenys,zodis[i].c_str());

            outf << A[0].duomenys << "\t" <<"\t";
            outf << A[1].duomenys << "\n";
            Edit5->Clear();
            Edit6->Clear();

        }
    outf.close();

}
//-----
// Potvyniu metu uzlietu plotu duomenu bazes mygtukas "Rodyti"
//-----
void __fastcall TForm1::Button4Click(TObject *Sender)
{

```



```

    outf <<"Sutampantys metai "<<A[0].duomenys <<" ir " <<A[1].duomenys << endl;
    outf <<"k" << "\t" <<"Paskaiciuotos tikimybes:" << endl;
        for (i=1;i<=100;i++){
            rk[i]=1-pow(funk,i);
            Form2->Memol->Lines-
>Append(IntToStr(i)+"\t\t"+FloatToStrF(rk[i],ffFixed,7,5));
            strcpy(A[1].duomenys,FloatToStrF(rk[i],ffFixed,7,5).c_str());
            strcpy(A[0].duomenys,IntToStr(i).c_str());
            outf <<A[0].duomenys<<"\t"<<"\t"<<A[1].duomenys<< endl;
        }
    outf << "
    Form2->Caption="Pasikartojimo tikimybes";
    Form2->ShowModal();
    Edit7->Clear();
    }
    outf.close();
}

}
//-----
// Pasikartojimu tikimybiu skaiciavimo mygtukas "Visi skaiciavimai"
//-----
void __fastcall TForm1::Button6Click(TObject *Sender)
{
    char *failas="rk.dat"; //duomeniu failo pavadinimas
    Form2->Caption="Pasikartojimu tikimybes";
    Form2->Memol->Lines->LoadFromFile(failas);
    Form2->ShowModal();
}
//-----
// Pervirsiu tikimybiu skaiciavimo mygtukas "Skaiciuoti"
//-----
void __fastcall TForm1::Button7Click(TObject *Sender)
{
    int nr,i,j;
    AnsiString zodis[4];
    int sk,poz=0,r,jj,k;
    float lygis[4];
    double p;
    char eil[30],*rod;
    double t1,t2;
//-----
zodis[0]=Edit8->Text; //metai r
zodis[1]=Edit9->Text; //kartai jj
zodis[2]=Edit10->Text; //dydis k

nr=skaitymas1();
//-----
//ar duomenys ivesti teisingai?
//-----
poz=0;
strcpy(eil,zodis[0].c_str());
lygis[0]=strtod(eil,&rod);
if (atoi(eil)> nr-2 || *rod!=NULL || lygis[0]==0){
    Application->MessageBox("Neteisingai ivesti
metai!", "Pranesimas",MB_OK+MB_ICONERROR);
    poz=1;
}
else r=atoi(eil);

strcpy(eil,zodis[1].c_str());
lygis[1]=strtod(eil,&rod);
if (atoi(eil)>r || *rod!=NULL){
    Application->MessageBox("Neteisingai ivestas dydis
j!", "Pranesimas",MB_OK+MB_ICONERROR);
    poz=1;
}
else jj=atoi(eil);

strcpy(eil,zodis[2].c_str());
lygis[2]=strtod(eil,&rod);

```

```

if (atoi(eil)> nr-2 || *rod!=NULL ||lygis[0]==0 ){
    Application->MessageBox("Neteisingai ivetas dydis
k!", "Pranesimas", MB_OK+MB_ICONERROR);
    poz=1;
}
else k=atoi(eil);
//-----
if (poz==0) {
//-----
// Rikiavimas max->min
//-----
for(i=3;i<nr;i++){
for (j=3;j<nr;j++){
if (B[j].smal_lyg<B[j+1].smal_lyg){
    B[0].smal_lyg=B[j].smal_lyg;
    B[j].smal_lyg=B[j+1].smal_lyg;
    B[j+1].smal_lyg=B[0].smal_lyg;
}
if (B[j].tilze_lyg<B[j+1].tilze_lyg){
    B[0].tilze_lyg=B[j].tilze_lyg;
    B[j].tilze_lyg=B[j+1].tilze_lyg;
    B[j+1].tilze_lyg=B[0].tilze_lyg;
}
if (B[j].rusne_lyg<B[j+1].rusne_lyg){
    B[0].rusne_lyg=B[j].rusne_lyg;
    B[j].rusne_lyg=B[j+1].rusne_lyg;
    B[j+1].rusne_lyg=B[0].rusne_lyg;
}
}
}
//-----
// Binomu skaiciavimas
//-----
nr=nr-2;
t1=faktorialas(r+nr-k-jj)*faktorialas(jj+k-1)*faktorialas(nr)*faktorialas(r);
t2=faktorialas(nr-k)*faktorialas(r-jj)*faktorialas(k-
1)*faktorialas(jj)*faktorialas(r+nr);
//p=t2;
p=t1*1.0/t2*1.0;
//p=10.0/3.0;
Form2->Memol->Lines->Clear();
Form2->Memol->Lines->Append(IntToStr(r)+" metu laikotarpyje Smalininku vandens lygis "+
FloatToStrF(B[k+3].smal_lyg,ffFixed,5,3)+" virsijamas
"+IntToStr(jj)+
" kartu su tikimybe" );
Form2->Memol->Lines->Append("p= "+FloatToStrF(p,ffFixed,7,5));
Form2->Memol->Lines->Append(IntToStr(r)+" metu laikotarpyje Tilzes vandens lygis "+
FloatToStrF(B[k+3].tilze_lyg,ffFixed,5,3)+" virsijamas
"+IntToStr(jj)+
" kartu su tikimybe" );
Form2->Memol->Lines->Append("p= "+FloatToStrF(p,ffFixed,7,5));

Form2->Memol->Lines->Append(IntToStr(r)+" metu laikotarpyje Rusnes vandens lygis "+
FloatToStrF(B[k+3].rusne_lyg,ffFixed,5,3)+" virsijamas
"+IntToStr(jj)+
" kartu su tikimybe" );
Form2->Memol->Lines->Append("p= "+FloatToStrF(p,ffFixed,7,5));

Form2->Caption="Pervirsiu tikimybes";
Form2->ShowModal();
Edit8->Clear();
Edit9->Clear();
Edit10->Clear();
}
}

//*****
// Meniu juostos aptranavimo funkcijos
//*****
//-----
// Meniu->Failas->Baigti darba

```

```

//-----
void __fastcall TForm1::BaigtisClick(TObject *Sender)
{
    Close();
}
//-----
// Meniu->Pagalba->Aprasymas
//-----
void __fastcall TForm1::Aprasymas1Click(TObject *Sender)
{
    char *failas="Pagalba.hlp"; //duomenu failo pavadinimas
    Form3->Memo1->Lines->LoadFromFile(failas);
    Form3->ShowModal();
}
//-----
// Meniu->Failas->Apie programa
//-----
void __fastcall TForm1::Apieprograma1Click(TObject *Sender)
{
    Form4->ShowModal();
}
//*****
// Paprogrames
//*****
//-----
// Faktorialo skaiciavimo paprograme
//-----
double TForm1::faktorialas(int sk)
{
    int i;
    double rez=1;
    for(i=1;i<=sk;i++){
        rez=rez*i;
    }
    return(rez);
}
//-----
// Potvyniu lygiu duomenu nuskaitymo paprograme
//-----
int TForm1::skaitymas1() //lygiu failo nuskaitymas
{
    FILE *fv;
    const sk=300;
    char *failas="data.dat";
    char buf[sk],R[200];
    int n,i=0;
    n=0;
    fv=fopen("data.dat","r");
    ifstream inf (failas, ios:: in);
    if (!inf){Application->MessageBox("Neatidarytas
failas!", "Pranesimas", MB_OK+MB_ICONERROR);}
    else {
        while (!feof(fv)){fgets(R,80,fv);n++;} //suskaiciuojamos duomenu failo eilutes
        i=0;
        B=new sarasas2[n];
        if (n-1==3){
            i=2;
            B[i].tilze_R=1;
            B[i].smal_R=1;
            B[i].rusne_R=1;
        }
        else {
            i=0;
            while (inf.getline(buf,sk)) //nuskaitytos eilutes iskaidymas
            {
                if (i>=3){
                    B[i].metai=atoi(strtok(buf,"\t"));
                    B[i].smal_lyg=skaicius(strtok(NULL,"\t"));
                    B[i].smal_R=skaicius(strtok(NULL,"\t"));
                    B[i].tilze_lyg=skaicius(strtok(NULL,"\t"));
                    B[i].tilze_R=skaicius(strtok(NULL,"\t"));
                    B[i].rusne_lyg=skaicius(strtok(NULL,"\t"));
                }
            }
        }
    }
}

```

```

        B[i].rusne_R=skaicius(strtok(NULL, " "));
    }
    i++;
}
i--;
}

inf.close();
return(i);
}
}
//-----
// Potvyniu metu uzlietu plotu duomenu bazes nuskaitymas
//-----
int TForm1::skaitymas2()          //plotu failo nuskaitymas
{
    FILE *fv;
    const sk=300;
    char *failas="plotas.dat";
    char buf[sk],R[200];
    int n,i=0;
    n=0;
    fv=fopen("plotas.dat","r");
    ifstream inf (failas, ios:: in);
    if (!inf){Application->MessageBox("Neatidarytas
failas!","Pranesimas",MB_OK+MB_ICONERROR);}
    while (!feof(fv)){fgets(R,80,fv);n++;} //suskaiciuojamos duomenu failo eilutes
    i=0;
    C=new sarasas3[n];
    //-----;
    i=0;
    while (inf.getline(buf,sk) //nuskaitytos eilutes iskaidymas
    {
        if (i>=2){          //atmetami lentutes stulpeliu aprasymai
            C[i].metai=atoi(strtok(buf,"\t"));
            C[i].s=skaicius(strtok(NULL,"\n"));
        }
        i++;
    }
    i--;

    //Mem1->Lines->Append(i);

    inf.close();
    return(i);
}

//-----
// Masyvo maksimalios reiksmes ieskojimo paprograme
//-----
void TForm1::maksimumas(int j)
{
    int i;
    m[0]=B[3].smal_lyg;
    for (i=4;i<=j;i++)
    {
        if (B[i].smal_lyg>m[0]) m[0]=B[i].smal_lyg;
    }

    m[1]=B[3].tilze_lyg;
    for (i=4;i<=j;i++)
    {
        if (B[i].tilze_lyg>m[1]) m[1]=B[i].tilze_lyg;
    }

    m[2]=B[3].rusne_lyg;
    for (i=4;i<=j;i++)
    {
        if (B[i].rusne_lyg>m[2]) m[2]=B[i].rusne_lyg;
    }
}

```

```

    }
//-----
// Nuskaitytos duomenų eilutės konvertavimas į skaičių
//-----
float TForm1::skaicius(char *mas)
{int k,c,t;
  char *zime=",";
  char *rod;
  float L;
  k=strcspn(mas,zime); //ieskomas ilgis iki kablelio
  c=strlen(mas); //visas eilutės ilgis
  if (c!=k){
    mas[k]='.';
  }
  L=strtod(mas,&rod);
  return(L);
}

```

Modulių sąsajos:

TFORM1

```

//-----

#ifdef PotviniaiH
#define PotviniaiH
//-----
#include <Classes.hpp>
#include <Controls.hpp>
#include <StdCtrls.hpp>
#include <Forms.hpp>
#include <ComCtrls.hpp>
#include <ExtCtrls.hpp>
#include <Menus.hpp>
//-----
struct sarasas
{ char duomenys[30];
};
struct sarasas2
{ int metai;
  float smal_lyg;
  int smal_R;
  float tilze_lyg;
  int tilze_R;
  float rusne_lyg;
  int rusne_R;
};
struct sarasas3
{ int metai;
  float s;
};

class TForm1 : public TForm
{
__published: // IDE-managed Components
  TPageControl *PageControl1;
  TTabSheet *TabSheet1;
  TTabSheet *TabSheet2;
  TEdit *Edit1;
  TEdit *Edit2;
  TEdit *Edit3;
  TEdit *Edit4;
  TButton *Button1;
  TButton *Button2;
  TEdit *Edit5;
  TEdit *Edit6;
  TButton *Button3;
  TButton *Button4;
  TEdit *Edit7;
  TButton *Button5;
  TButton *Button6;

```

```

TEdit *Edit8;
TEdit *Edit9;
TEdit *Edit10;
TButton *Button7;
TLabel *Label3;
TLabel *Label11;
TLabel *Label12;
TLabel *Label7;
TLabel *Label4;
TLabel *Label5;
TLabel *Label6;
TLabel *Label17;
TLabel *Label18;
TLabel *Label19;
TLabel *Label20;
TBevel *Bevel1;
TLabel *Label8;
TLabel *Label9;
TLabel *Label10;
TLabel *Label11;
TBevel *Bevel2;
TLabel *Label15;
TLabel *Label16;
TMainMenu *MainMenu1;
TMenuItem *Failas1;
TMenuItem *Pagalba1;
TMenuItem *Baigti1;
TMenuItem *Aprasymas1;
TMenuItem *Apieprograma1;
void __fastcall Button1Click(TObject *Sender);
void __fastcall Button2Click(TObject *Sender);
void __fastcall Button3Click(TObject *Sender);
void __fastcall Button4Click(TObject *Sender);
void __fastcall Button5Click(TObject *Sender);
void __fastcall Button6Click(TObject *Sender);
void __fastcall Button7Click(TObject *Sender);
void __fastcall Baigti1Click(TObject *Sender);
void __fastcall Aprasymas1Click(TObject *Sender);
void __fastcall Apieprograma1Click(TObject *Sender);
private: // User declarations
sarasas A[4];
sarasas2 *B;
sarasas3 *C;
int skaitymas1();
int skaitymas2();
double faktorialas(int);
float skaicius(char*);
void maksimumas(int);
public: // User declarations
__fastcall TForm1(TComponent* Owner);
};
//-----
extern PACKAGE TForm1 *Form1;
//-----
#endif

TFORM2
//-----

#ifndef IsvedimasH
#define IsvedimasH
//-----
#include <Classes.hpp>
#include <Controls.hpp>
#include <StdCtrls.hpp>
#include <Forms.hpp>
//-----
class TForm2 : public TForm
{
__published: // IDE-managed Components

```

```

        TMemo *Memol;
        void __fastcall MemolChange(TObject *Sender);
private:    // User declarations
public:    // User declarations
        __fastcall TForm2(TComponent* Owner);
};
//-----
extern PACKAGE TForm2 *Form2;
//-----
#endif

TForm3
//-----

#ifndef apieH
#define apieH
//-----
#include <Classes.hpp>
#include <Controls.hpp>
#include <StdCtrls.hpp>
#include <Forms.hpp>
#include <Mask.hpp>
#include <ExtCtrls.hpp>
#include <jpeg.hpp>
//-----
class TForm4 : public TForm
{
__published: // IDE-managed Components
        TLabel *Label1;
        TButton *Button1;
        TStaticText *StaticText1;
        TStaticText *StaticText2;
        TStaticText *StaticText3;
        TImage *Image1;
        TStaticText *StaticText4;
        TLabel *Label2;
        TLabel *Label3;
        TLabel *Label4;
        TLabel *Label5;
        TLabel *Label6;
        void __fastcall Button1Click(TObject *Sender);
private:    // User declarations
public:    // User declarations
        __fastcall TForm4(TComponent* Owner);
};
//-----
extern PACKAGE TForm4 *Form4;
//-----
#endif

```

5.2 LAIKO EILUČIŲ MODELIŲ SUDARYMAS

```

data lygiai;
input metai smalininkai rusne; /*Kintamuju ivedimas*/
cards;
1961 11.29 0.96
1962 13.94 2.66
1963 12.99 1.55
1964 13.25 1.46
1965 12.71 1.34
1966 11.99 1.03
1967 11.85 1.15
1968 13.46 1.96
1969 11.74 1.94
1970 14.73 2.47
1971 12.47 1.31
1972 10.40 0.59

```

```

1973 10.97 0.72
1974 9.76 0.08
1975 11.41 0.71
1976 12.16 0.99
1977 11.60 0.84
1978 11.94 0.94
1979 14.61 2.67
1980 12.94 2.27
1981 11.83 1.07
1982 11.70 1.17
1983 12.79 1.51
1984 11.31 0.81
1985 13.20 1.95
1986 13.20 1.77
1987 13.93 2.49
1988 13.48 1.97
1989 10.81 0.83
1990 9.67 0.77
1991 10.67 0.73
1992 10.31 0.77
1993 11.16 1.05
1994 13.73 2.10
1995 11.54 1.27
1996 13.20 1.99
1997 9.77 0.79
1998 11.06 1.05
1999 12.53 1.40

```

```
;
```

```
run;
```

```
proc univariate data=lygiai;
```

```
var smalininkai tilze rusne;
```

```
run;
```

```

proc arima data=lygiai; /*Procedura ARIMA kintamajam rusne*/
identify var=rusne scan esacf alpha=0.05; /*modelio eiles nustatymas*/
identify var=rusne; /*Modelio MA(1) vertinimas*/
estimate p=1 q=0 method=uls;
identify var=rusne; /*Modelio ARIMA(1,1,0) vertinimas*/
estimate p=0 q=4 method=uls;
identify var=rusne; /*Modelio ARIMA(3,2,0) vertinimas*/
estimate p=2 q=0 method=uls;
identify var=rusne(1); /*Modelio ARIMA(2,2,1) vertinimas*/
estimate p=4 q=1 method=uls;
run;

```

```

proc arima data=lygiai; /*Procedura ARIMA kintamajam smalininkai*/
identify var=smalininkai scan esacf alpha=0.05; /*modelio eiles nustatymas*/
identify var=smalininkai; /*Modelio ARIMA(1,1,1) vertinimas*/
estimate p=1 q=0 method=uls;*/
identify var=smalininkai; /*Modelio ARIMA(2,1,0) vertinimas*/
estimate p=0 q=4 method=uls;*/
identify var=smalininkai(1); /*Modelio ARIMA(3,2,0) vertinimas*/
estimate p=4 q=0 method=uls;
run;

```

```
quit;
```


6 PRIEDAS. STRAIPSNIAI

6.1 STATISTINIŲ METODŲ TAIKYMAS EKSTREMALIESIEMS ĮVYKIAMS – POTVYNIAMS NEMUNO DELTOJE – TIRTI

Dr. Valerija KARPICKAITĖ, Kauno technologijos universiteto Fundamentaliųjų mokslų fakulteto Matematinės sistematikos katedros docentė.

Pagrindinės mokslinių tyrimų kryptys – matematika, informatika, fiziniai mokslai.

Adresas: KTU, Studentų g. 50, Kaunas. Tel. (8-37) 45-16-58.

Julija SMIRNOVA, Kauno technologijos universiteto magistrantė.

Pagrindinės mokslinių tyrimų kryptys – matematika, ekstremalieji įvykiai.

Adresas: KTU, Studentų g. 50, Kaunas. El. p. julija.sm@one.lt.

SANTRAUKA

Kiekvienais metais per pavasario potvynius Nemuno deltoje užliejamas iki 52,4 tūkst. ha plotas. Potvynių padariniai veikia žmonių gyvenimo kokybę ir ūkinę veiklą.

Pagrindinis šio darbo tikslas – susipažinti su užsienyje taikomais statistiniais ekstremaliųjų įvykių tyrimo metodais: Gumbelo paviršių metodu[2], pasikartojimo periodo metodu[2] ir pritaikyti juos Lietuvai būdingam ekstremaliajam įvykiui – potvyniui: apskaičiuoti paviršių tikimybes ir sudaryti regresijos modelį, kuris padėtų prognozuoti užliejamą plotą pagal Tilžės ir Smalininkų postuose matuojamus vandens lygius.

Prasminiai žodžiai: *potvyniai, tikimybės, veiksniai, regresijos modelis.*

IVADAS

Potvynius Nemuno deltoje galime priskirti prie ekstremaliųjų įvykių. Jie kelia nemažą problemų užliejamų plotų gyventojams, civilinės saugos tarnyboms, veikia žmonių ūkinę veiklą ir pasunkina gyvenimo sąlygas palyginti didelėje teritorijoje.

Sovietmečiu buvo tiriami potvyniai ir vertinami jų padaryti nuostoliai, bet vėliau šie darbai nutrūko. Dabar jie vėl atnaujinti Lietuvos žemės ūkio universiteto Vandens ūkio institute.

Naudojant statistinius ekstremaliųjų įvykių metodus, taikomus užsienyje, darbe bandoma vertinti kai kurias potvynių charakteristikas.

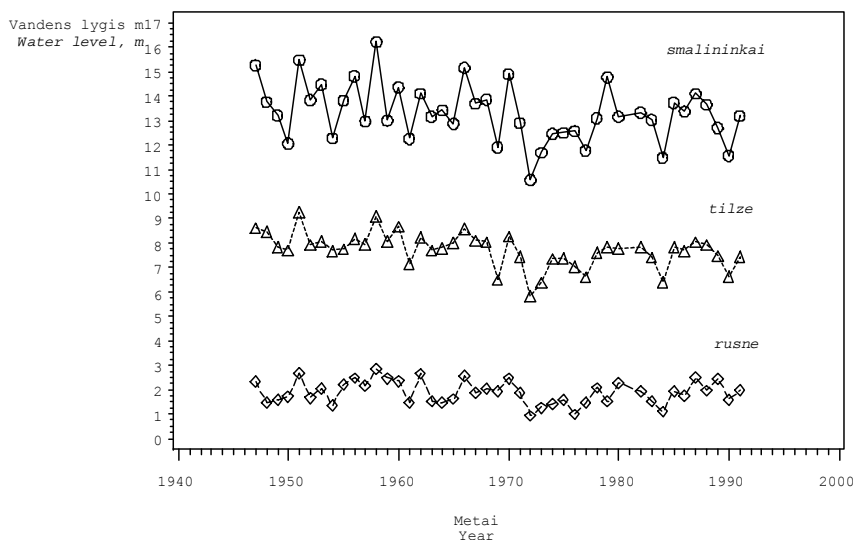
DUOMENŲ PARENGIMAS ANALIZEI

Nors potvynių stebėjimų imtį sudaro daug stebėjimų kiekvienais metais, bet analizei dažniausiai naudojama tik didžiausia metinė reikšmė. Taip išvengiama vidinės metų priklausomybės. 1 paveiksle pavaizduoti didžiausi metiniai vandens lygiai Smalininkų, Tilžės bei Rusnės vandens matavimo postuose. Čia ir toliau bus naudojami tokie kintamieji:

smalininkai – Smalininkų vandens matavimo posto vandens lygių kintamasis;

tilze – Tilžės vandens matavimo posto vandens lygių kintamasis;

rusne – Rusnės vandens matavimo posto vandens lygių kintamasis.



1 pav. Didžiausi metiniai vandens lygiai Smalininkų, Tilžės bei Rusnės vandens matavimo postuose

Fig. 1. Maximum annual water levels at Smalininkai, Tilže and Rusne measurement stations

Kaip matyti iš grafiko, duomenų sklaidų pobūdis yra vienodas, skiriasi tik atstumas iki abscisės. Taip esti todėl, kad šiuose matavimo postuose skirtingos nulio altitudės: Smalininkų poste – 7,38 m; Tilžės – 2,06 m; Rusnės – 1,53 m.

Imčių statistinė analizė atliekama su paketu SAS [5]. Visų trijų kintamųjų imtys turi tokias statistikas.

1 lentelė. Kintamųjų *smalininkai*, *tilze* ir *rusne* statistikos

Table 1. Statistics for *smalininkai*, *tilze* and *rusne* variables

	<i>smalininkai</i>	<i>tilze</i>	<i>rusne</i>
Imties dydis n / Data-set size	44	44	44
Mažiausia reikšmė / Minimum	10,580	5,84	0,94
1 kvartilis / 1st quartile	12,535	7,410	1,530
Mediana / Median	13,205	7,810	1,91000
Vidurkis / Mean	13,33295	7,71477	1,89341
3 kvartilis / 3rd quartile	13,980	8,065	2,295
Didžiausia reikšmė / Maximum	16,220	9,250	2,850
Dispersija / Dispersion	1,39827	0,49465	0,22981

DIDŽIAUSIO VANDENS LYGIO PERVIRŠIO TIKIMYBĖS

Tegul $X_{n,n} < \dots < X_{1,n}$ - sutvarkyta imties X_1, \dots, X_n statistika, įstatyta į begalinę nepriklausomų vienodų skirstinių seką (X_i) su tolydžia pasiskirstymo funkcija F . k -tos viršutinės eilės statistika $X_{k,n}$ imama kaip atsitiktinio slenksčio reikšmė ir žymima $S_r^n(k)$, čia $r \geq 1$ - $X_{k,n}$ perviršių skaičius tarp sekančių r stebėjimų X_{n+1}, \dots, X_{n+r} , t.y.

$$S_r^n(k) = \sum_{i=1}^r I_{\{X_{n+i} > X_{k,n}\}}.$$

Paprastumo dėlei $S_r^n(k)$ žymėsime S .

Atsitiktinis dydis S turi hypergeometrinį skirstinį, t.y

$$P(S = j) = \frac{\binom{r+n-k-j}{n-k} \binom{j+k-1}{k-1}}{\binom{r+n}{n}}, j = 0, 1, \dots, r. - \text{tikimybė, kad } k\text{-toji pagal didumą per stebėjimo}$$

laikotarpį vandens lygio reikšmė bus didesnė j kartų per ateinančius r (šiuo atveju $r = 10$) metų.

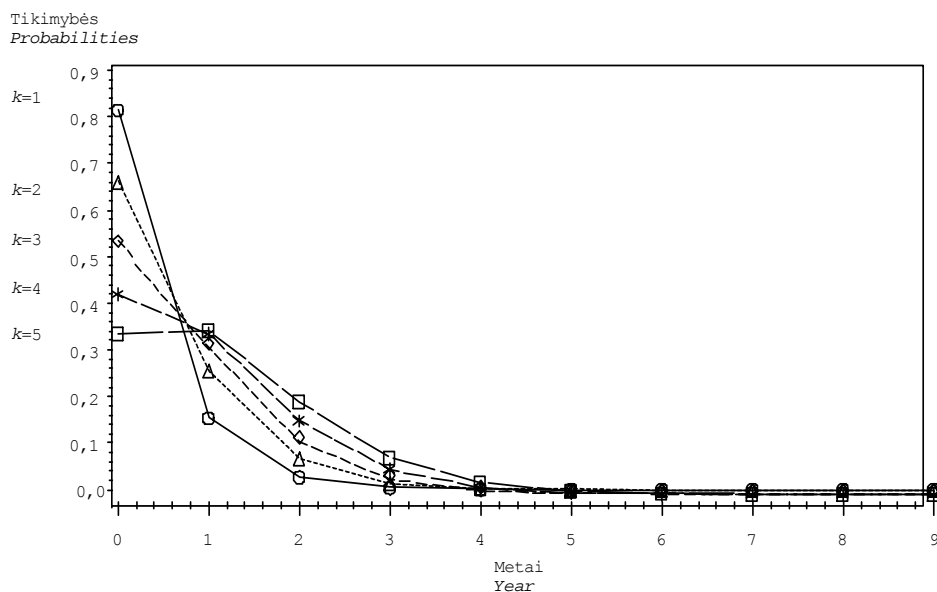
Šis metodas vadinasi Gumbelo perviršio metodu [2].

2 lentelė. Didžiausių vandens lygių perviršių tikimybės ateinantiems 10 metų

Table 2. Exceedance probabilities for forthcoming 10 years

P	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
$j = 0$	0,81481	0,66108	0,53395	0,42925	0,34340
$j = 1$	0,15374	0,25426	0,31409	0,34340	0,35041
$j = 2$	0,02661	0,06730	0,11307	0,15768	0,19710
$j = 3$	0,00417	0,01436	0,03077	0,05256	0,07828
$j = 4$	0,00058	0,00256	0,00673	0,01370	0,02383
$j = 5$	0,00007	0,00038	0,00120	0,00286	0,00572
$j = 6$	0,00001	0,00005	0,00017	0,00048	0,00108
$j = 7$	0	0	0,00002	0,00006	0,00016
$j = 8$	0	0	0	0,00001	0,00002
$j = 9$	0	0	0	0	0

2 lentelėje pateiktos tikimybės pavaizduotos 2 paveiksle.



2 pav. Perviršio tikimybės atitinkamoms k reikšmėms

Fig 2. Exceedance probabilities for corresponding k values

Remiantis gautais rezultatais galima padaryti išvadą: tikimybės, kad per ateinančius 10 metų vandens lygiai stebėjimo stotyse (postuose) nebus didesni nei nurodytieji 3 lentelėje, yra tokios:

3 lentelė. Atitinkamos tikimybės vandens lygių ribos ribos per ateinančius 10 metų
 Table 3. Level that will not be exceeded during forthcoming 10 years, with calculated probabilities

Tikimybė Probability	Vandens lygis, kuris nebus didesnis, m Water level that will not be exceeded, m		
	Smalininkų v. m. p.	Tilžės v. m. p.	Rusnės v. m. p.
0,81481	16,22	9,25	2,85
0,66108	15,48	9,09	2,68
0,53395	15,26	8,66	2,66
0,42925	15,16	8,61	2,58
0,34340	14,90	8,59	2,49

PASIKARTOJIMO PERIODAS

Pasikartojimo periodo metodikai [2] pritaikyti reikalingi duomenys apie didžiausius Šilutės rajono potvynius (4 lentelė).

4 lentelė. Didžiausių Šilutės rajono potvynių suvestinė
 (Šilutės rajono civilinės ir priešgaisrinės saugos tarnybos duomenys)

Table 4. Summary of the largest floods in Silutė district (according to the date of civil and fire service of Silutė district)

Metai Year	Užtvindyta teritorija ha Inundated territory, ha
1958	42 000
1975	48 000
1978	47 000
1982	45 000
1994	48 000
1999	45 000
2002	23 000

Kaip matyti iš suvestinės, per 1975 ir 1994 m. potvynius buvo užtvindytos vienodai didelės teritorijos – 48 000 ha. Laiko tarpas tarp šių įvykių – 19 metų. Taigi galima apibrėžti “19 metų” įvykį – “ypatingai didelį potvynį” taip: potvynis, kurio metu užtvindoma 48 000 ha teritorija.

Pasikartojimo periodo metode [2] pirmo pasisekimo laikas

$$L(u) = \min\{i \geq 1 : X_i > u\},$$

t.y. laikas pirmo perviršio per slenkstį u , yra geometrinis atsitiktinis dydis su skirstiniu

$$P(L(u) = k) = (1 - p)^{k-1} p, k = 1, 2, \dots$$

Pagal apibrėžimą $EL(u) = \frac{1}{p} = 19$, iš čia išeina, kad pasisekimo tikimybė lygi $p = \frac{1}{19} = 0,05263$.

Taigi tikimybė, kad per k stebėjimų (t. y. per k metų) bent vieną kartą potvynio užtvindyta teritorija sudarys 48 000 ha, apskaičiuojama taip:

$$r_k = p \sum_{i=1}^k (1 - p)^{i-1} = 1 - (1 - p)^k = 1 - (1 - 0.05263)^k, \quad k \in \mathbb{N}.$$

Apskaičiuotos tikimybės pateikiamos 5 lentelėje:

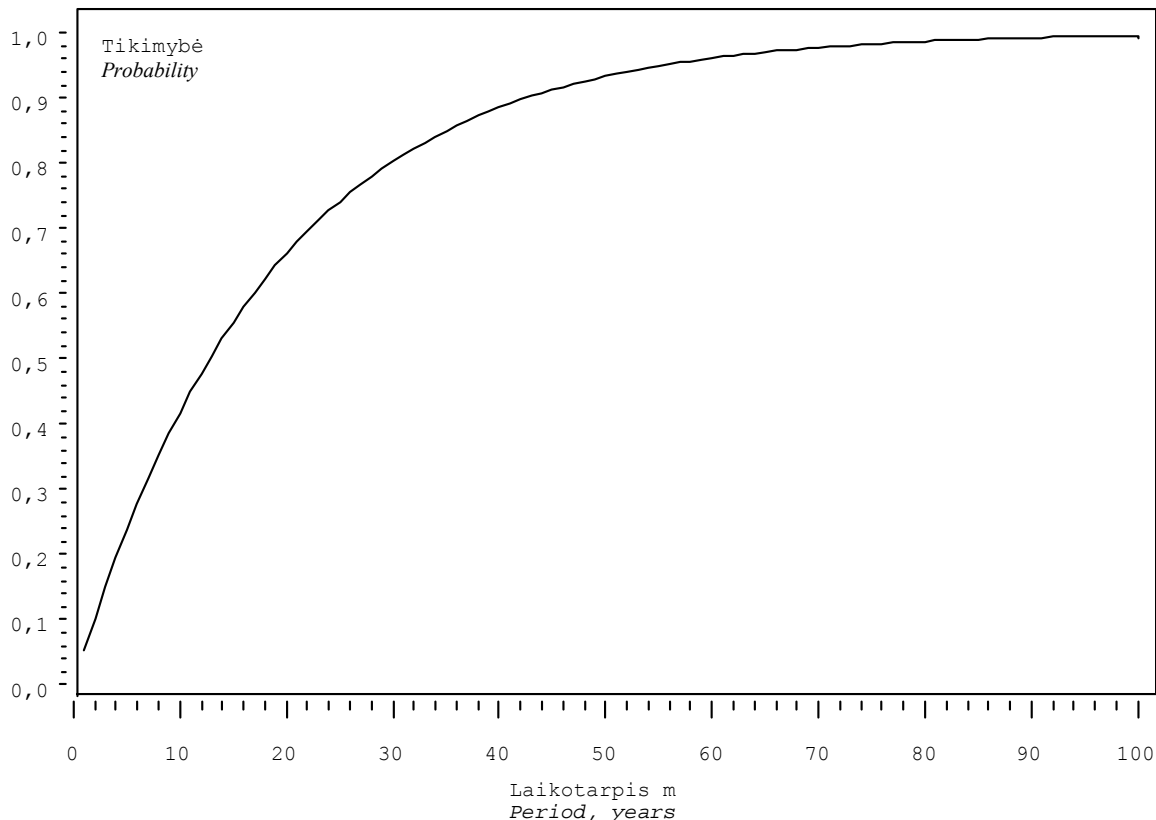
5 lentelė. Tikimybės, kad bus “ypatingai didelis potvynis”, k metų laikotarpiu

Table 5. Probability of a t “particularly large flood” in k - year period

k	r_k
1	0,05263
2	0,10249
3	0,14973

4	0,19448
5	0,23687
10	0,41763
20	0,66085
50	0,93301
100	0,99551

5 lentelėje pateiktos tikimybės pavaizduotos 3 paveiksle.



3 pav. Tikimybė, kad per stebėtą laikotarpį įvyks bent vienas “ypatingai didelis potvynis”

Fig 3. Probability of at least one “particularly large flood” during the observed period

VEIKSNIAI, TURINTYS ĮTAKOS POTVYNIAMS

Nustatyta [3, 4], kad potvynių dydžiui ir daromai žalai turi įtakos tokie veiksniai:

- upių vandens lygiai (kad išvengtume vidinės metų priklausomybės, imama tik didžiausia metinė reikšmė);
- ledų sangrūdos;
- išalo poveikis apsauginių pylimų išplovimui;
- potvynio trukmė;
- užtvindoma teritorija;
- gelbėjimo darbų organizavimas;
- kiti veiksniai.

Darbe dėl duomenų trūkumo nagrinėjamas supaprastintas modelis. Duomenys gauti iš dviejų šaltinių: LŽŪU doc. A.Dumbrasuko ir Šilutės rajono civilinės ir priešgaisrinės saugos tarnybos.

6 lentelė. Potvynių metu Šilutės rajonui padaryti nuostoliai ir užtvindyti plotai
Table 6. The largest inundated territory and loss during floods in Silutė district

Metai Year	Maksimalus vandens lygis matavimo postuose, m The highest water levels at the measurement points, m			Užtvindyta teritorija ha Inundated territory, ha	Nuostoliai*, tūkst. Lt Loss, thousands Lt
	Tilžė	Rusnė	Smalininkai		
1958	9,09	2,85	16,22	42 000	-
1962	8,25	2,66	14,11	-	518,18
1963	7,69	1,55	13,16	-	291,5
1964	7,80	1,46	13,42	-	205
1965	8,00	1,62	12,88	-	462,2
1966	8,59	2,58	15,16	-	659,1
1967	8,09	1,87	13,70	-	155,25
1968	1968	2,04	13,86	-	695
1969	6,49	1,94	11,91	-	418,6
1975	7,36	1,60	12,50	48 000	-
1978	7,62	1,47	13,10	47 000	-
1982	7,84	1,93	13,33	45 000	-
1994	-	-	-	48 000	1 400
1999	-	-	-	45 000	1 251,93
2002	-	-	-	23 000	554,29

* 1962 – 1969 metų patirti nuostoliai rubliais yra perskaičiuoti, remiantis 1999 metų apytikriu lito perkamosios galios koeficientu $1 \text{ rb} \cong 8 \text{ Lt}$

Dėl skirtingų duomenų šaltinių ir jų neatitikimo neįmanoma sudaryti regresijos modelio:

$$\text{nuostoliai} = a + b_1 \cdot \text{tilze} + b_2 \cdot \text{rusne} + b_3 \cdot \text{smalininkai} + b_4 \cdot \text{plotas},$$

todėl bus nagrinėjamas toks modelis:

$$\text{plotas} = a + b_1 \cdot \text{tilze} + b_2 \cdot \text{rusne} + b_3 \cdot \text{smalininkai}.$$

REGRESIJOS MODELIO SUDARYMAS

Patikrinome, ar egzistuoja užtvindomo ploto priklausomybė nuo didžiausių vandens lygių Tilžės, Rusnės bei Smalininkų vandens matavimo postuose. Iš žemiau pateiktų grafikų matyti, kad priklausomybė yra tiesinė:

Pritaikėme daugialypės regresijos metodą [1] ir sudarėme tokį regresijos modelį:

$$\text{plotas} = a + b_1 \cdot \text{tilze} + b_2 \cdot \text{rusne} + b_3 \cdot \text{smalininkai},$$

čia *tilze*, *rusne*, *smalininkai* – maksimalių vandens lygių atitinkamuose matavimo postuose kintamieji;

plotas – užtvindyto ploto kintamasis.

Visiems reikalingiems koeficientams apskaičiuoti ir regresijos ryšiui nustatyti panaudota SAS paketo procedūra REG [5]. Gauname, kad kintamasis *rusne* nėra įtrauktas į regresijos modelį, nes ryšys tarp kintamųjų *plotas* ir *rusne* nėra statistiškai reikšmingas:

7 lentelė. SAS procedūros REG rezultatai (SAS Output)

Table 7. Results of procedure REG (SAS Output)

Variable	Parameter		Standard		
	Estimate	Error	Type II SS	F Value	Pr > F
Intercept	91862	4873.67187	23255451	355.27	0.0337
tilze	-15741	2918.40253	1904336	29.09	0.1167
smalininkai	5745.23046	1349.57798	1186273	18.12	0.1469

 Bounds on condition number: 229.74, 918.96

All variables left in the model are significant at the 0.1500 level.
No other variable met the 0.1500 significance level for entry into the model.

Gautas regresijos modelis:

$$plotas = 91862 - 15741 \cdot tilze + 5745,23 \cdot smalininkai.$$

Sudarytas modelis, ko gero, nėra visiškai tikslus hidrologiniu požiūriu, nes kintamojo *tilze* koeficientas rodo, kad užliejamas plotas didėja mažėjant vandens lygiui Tilžės vandens matavimo poste. Toks neatitikimas galimas dėl pradinių statistikos duomenų trūkumo, ledų sangrūdų poveikio ar pan.. Maža pradinė imtis nulemia netikslius rezultatus. Taigi šį modelį reikėtų vertinti ne tiek hidrologiniu, kiek statistiniu požiūriu.

Patikrinę statistines modelio prielaidas įsitikiname, kad visos jos tenkinamos. Taigi sudarytas regresijos modelis yra pakankamai geras, nes koreguotas determinacijos (apibrėžtumo) koeficientas $r_{adj}^2 = 0,9906$, o tai reiškia, kad 99% sklaidos galima paaiškinti.

Ateityje, turint daugiau statistikos duomenų, galima sudaryti tikslesnį regresijos modelį ir panaudoti jį potvynių daromiems nuostoliams prognozuoti atsižvelgiant į daugelį anksčiau minėtų veiksnių.

IŠVADOS

Susipažinus su užsienyje taikomais ekstremaliųjų įvykių tyrimo metodais ir pritaikius juos Lietuvos sąlygomis Nemuno deltoje kylantiems potvyniams tirti, gauti tokie rezultatai:

- taikant Gumbelo perviršių metodą, buvo apskaičiuotos tikimybės, rodančios, kad per ateinančius 10 metų vandens lygiai nustatyta skaičių kartų bus didesni už iki šiol deltoje užfiksuotuosius nuo didžiausio iki penkto dydžio;
- taikant pasikartojimo periodo metodą, buvo apskaičiuotos tikimybės, rodančios, kad per daugelį metų gali pasikartoti aprašytas “ypatingai didelis potvynis”, kurio užliejamas plotas sudaro 48 000 ha;
- taikant įrašų metodą, buvo nustatyta, kad tiriamų imčių duomenys yra nepriklausomi ir vienodai pasiskirstę;
- sudarytas užliejamo ploto regresijos modelis.

Taikant modelį nustatyta, kad užliejamas plotas gali būti prognozuotas pagal užfiksuotus vandens lygius Tilžės ir Smalininkų vandens matavimo postuose. Norint tiksliau prognozuoti būsimus nuostolius, reikėtų daugiau statistikos duomenų.

LITERATŪRA

1. Čekanavičius V., Murauskas G., 2002. Statistika ir jos taikymai. 2. Vilnius, TEV.
2. Embrechts P., Kluppelberg C., Mikosch T., 1997. Modelling Extremals Events. Springer-Verlag, Berlin Heidelberg.. –
3. Vaikasas S, Rimkus A., 1997. Potvynio vandens lygių dinamikos Nemuno deltoje tyrimai. Vandens ūkio inžinerija. Mokslo darbai. 2 (24), 108-125.
4. Vaikasas S, Rimkus A., 2000. Potvynio daromos Nemuno deltoje žalos ir naudos vertinimas. Vandens ūkio inžinerija. Mokslo darbai. 9 (31), 86-97.
5. Elliot, Rebecca J. Learning SAS in the Computer Lab. Duxbury Press, 1995, 175 p.

APPLICATION OF STATISTICAL METHODS FOR EXTREMAL EVENTS – THE NEMUNAS DELTA FLOODS

Valerija Karpickaitė, Julija Smirnova

Kaunas University of Technology

SUMMARY

Delta of the river Nemunas covers an area from the outlet of the tributary Jura to the Curonian Lagoon in Klaipėda District. In fact, every year an area up to 52,4 thousand hectares is inundated during the spring floods. The destructive flood consequences adversely affect life, farming conditions and property of 3,2 thousand inhabitants of Rusne and Panemune and other 46 villages.

The main purpose of this paper was to find out what statistical methods can be used to analyse floods as extreme events and to calculate probabilities of exceedance of the highest recorded water level. Also a regression model was made up to predict flooding area, which was based on the highest water levels in measuring points of Tilze and Smalininkai.

Dr. Valerija Karpickaitė, lecturer at the Kaunas University of Technology.

Research interests – mathematics, informatics, physical science.

Address: Studentu 50, LT – 3031 Kaunas, Lithuania. Phone: +370 37 45 16 58.

Julija Smirnova, master student at the Kaunas Technological University.

Research interests – mathematics, extreme events.

Address: Studentu 50, LT – 3031 Kaunas, Lithuania. E-mail: julija.sm@one.lt

6.2 VANDENS LYGIO NEMUNO ŽEMUPIO DELTOJE POTVYNIŲ METU ĮVERTINIMAS

J. Smirnova, doc. V. Karpickaitė

Kauno technologijos universitetas

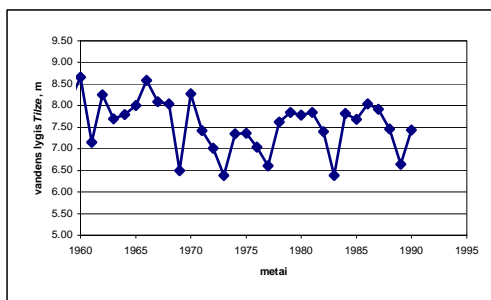
Gamtinės nelaimės, kurios atneša didžiulius nuostolius, laikomos ekstremaliais įvykiais. Įvertinus geografinę padėtį, Lietuvai negresia žemės drebėjimai – daugiausia nuostolių atnešančios gamtinės stichijos. Ekstremalių įvykių pavyzdžių Lietuvoje yra potvyniai Nemuno deltoje ir prie jos prigludusioje žemupio dalyje. 605 km² plote čia plyti vos vieną kitą metrą aukščiau jūros lygio pakilusi lyguma, kurią visą arba jos dalį kasmet, o kartais ir dažniau užlieja vanduo, kuris atneša didelius nuostolius žmonių turtui, o kartais ir gresia gyvybei.

Labai svarbiu faktoriumi laikomas potvynio metu kylančio vandens lygis. Kuo jis didesnis, tuo didesni plotai gali būti apsemti, ir tuo daugiau žalos gali būti patirta. Nuo seno vandens lygiai potvynio metu buvo stebimi ir fiksuojami, tačiau apie 1990 metus šis darbas nutrūko. Ir tik visai neseniai Kėdainių Vandens Ūkio Institutas jį atnaujino.

Hidrologiniams stebėjimams būdingi didesni ar mažesni svyravimai. Jeigu stebimi dydžiai išsidėsto vienas paskui kitą laiko atžvilgiu, tai gaunasi seka, kuri nagrinėjama kaip laiko eilutė. Pokyčių dėsningumas tokioje eilutėje turi tendencija mažėti didėjant laiko tarpui tarp stebėjimų. Kadangi šioms sekoms būdingas nereguliarumas, tai analizei geriausiai taikyti statistinius metodus.

Nagrinėjami hidrologiniai stebėjimai – skirtinguose Nemuno žemupio matavimo postuose užfiksuoti vandens lygio kintamieji. Analizė atliekama naudojant statistinio paketo SAS posistemę, kuri skirta laiko eilučių prognozavimui (SAS Time Series Forecasting System).

Nagrinėjamąjį kintamąjį *tilze*, sudaro 44 stebėjimai – 1947 – 1990 metų laikotarpio didžiausi metiniai vandens lygiai Tilžės matavimo poste, kurie pavaizduoti 1 grafike:



1 pav. Kintamojo *tilze* grafikas

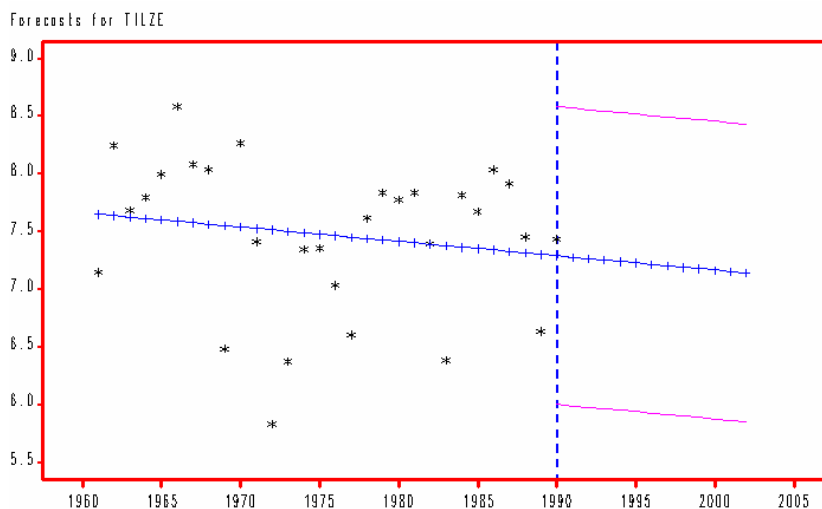
SAS Time Series Forecasting System Automatic Model Fit pagalba buvo sudaryti keli prognozavimo modeliai ir iš jų atrinkti 3 geriausiai atitinkantys. Modelio prognozės tikslumo matas – vidutinė kvadratinė paklaida pavaizduota 1 lentelėje.

1 lentelė

Atrinktų modelių statistiniai kriterijai

Statistinis kriterijus	Atrinkti modeliai		
	Tiesinis	Tiesinis eksponentinis išlyginamasis	Nuslopinto trendo eksponentinis išlyginamasis
Vidutinė kvadratinė paklaida	0,40372	0,41559	0,41575
Vidutinis absoliutus paklaidų procentas	7,16881	7,14301	7,13429
Vidutinė absoliutinė paklaida	0,51212	0,50504	0,50429

Iš lentelės matyti, kad sudarytas tiesinis modelis geriausias iš atrinktų. Šio modelio pagrindu sudaryta prognozė pavaizduota 2 pav. Jame taip pat parodyti viršutinis ir apatinis pasikliautino intervalo, kai $p = 0,95$ ($\alpha = 0,05$), rėžiai.



2 pav. Sudaryta tiesinė prognozė

Literatūra:

1. R. Rudzkis – Laiko eilučių analizė. Paskaitų konspektas, 2003.
2. Rebecca J. Elliott – Learning SAS in the Computer Lab, Duxbury Press, 1995.
3. Статистические методы в гидрологии – Гидрометеорологическое издательство, Ленинград, 1970.