

KAUNO TECHNOLOGIJOS UNIVERSITETAS
VYTAUTO DIDŽIOJO UNIVERSITETAS
LIETUVOS ENERGETIKOS INSTITUTAS

AISTĖ KARPUŠENKAITĖ

**PAVOJINGŲ IR KITŲ SPECIFINIŲ ATLIEKŲ SUSIDARYMO
PROGNOZAVIMO METODAI: LIETUVOS ATVEJIS**

Daktaro disertacijos santrauka
Technologijos mokslai, aplinkos inžinerija (T 004)

2019, Kaunas

Disertacija rengta 2014–2018 m. Kauno technologijos universiteto Cheminės technologijos fakulteto Aplinkos technologijų katedroje.

Mokslinis vadovas:

prof. dr. Gintaras DENAFAS (Kauno technologijos universitetas, technologijos mokslai, aplinkos inžinerija, T 004).

Mokslinis konsultantas:

doc. dr. Tomas RUZGAS (Kauno technologijos universitetas, Gamtos mokslai, matematika, N 001).

Redagavo: Aurelija Gražina Rukšaitė (Leidykla „Technologija“)

Aplinkos inžinerijos mokslo krypties disertacijos gynimo taryba:

Prof. habil. dr. Jurgis Kazimieras STANIŠKIS (Kauno technologijos universitetas, technologijos mokslai, aplinkos inžinerija, T 004) – **pirmininkas**;

Doc. dr. Dimitrios KOMILIS (Thrace Democritus universitetas, technologijos mokslai, aplinkos inžinerija, T 004);

Prof. dr. Mait KRIIPSALU (Estijos gyvybės mokslų universitetas, technologijos mokslai, aplinkos inžinerija, T 004).

Prof. dr. Jolita KRUOPIENĖ (Kauno technologijos universitetas, technologijos mokslai, aplinkos inžinerija, T 004);

Prof. dr. Arvydas POVILAITIS (Vytauto Didžiojo universitetas, technologijos mokslai, aplinkos inžinerija, T004).

Disertacija bus ginama viešame Aplinkos inžinerijos mokslo krypties disertacijos gynimo tarybos posėdyje 2019 m. rugpjūčio 30 d. 10 val. Kauno technologijos universiteto Rektorato salėje.

Adresas: K. Donelaičio g. 73-402, 44249 Kaunas, Lietuva.

Tel. (370) 37 30 00 42; faks. (370) 37 32 41 44; el. paštas doktorantura@ktu.lt.

Disertacijos santrauka išsiųsta 2019 m. liepos 15 d.

Su disertacija galima susipažinti internete <http://ktu.edu>, Kauno technologijos universiteto (K. Donelaičio g. 20, Kaunas), Vytauto Didžiojo universiteto (K. Donelaičio g. 52, Kaunas) bibliotekose ir Lietuvos energetikos institute (Breslaujos g. 3, Kaunas).

IŽANGA

Problematika – apžvelgus anksčiau vykdytų ir aprašytų mokslinių tyrimų rezultatus galima teigti, kad daugiausia matematinio modeliavimo metodų ir modelių buvo naudota prognozuojant komunalinių atliekų susidarymą, daug mažiau – atliekant pavojingų atliekų (praktiškai tik medicininių atliekų) susidarymo prognozes. Neišnaudota galimybė prognozuoti pavojingų ir kitų specifinių (ne komunalinių) atliekų susidarymą gali nulemti klaidingų arba blogai apskaičiuotų sprendimų priėmimą atliekų tvarkymo sektoriuje trumpuoju arba vidutiniu laikotarpiu. Žinant prognozuojamų atliekų kiekius ir / ar jų svyravimo tendencijas, galima atitinkamai planuoti ir realizuoti turimus ar būsimus atliekų tvarkymo sektoriaus pajėgumus. Tai leidžia išvengti dėl neefektyviai naudojamos sistemos ar jos neatitikimo realiai situacijai susidarantių ekonominių nuostolių.

Tyrimo objektas – medicininių atliekų, automobilinių atliekų, dienos šviesos lempų ir gyvsidabrio turinčių atliekų susidarymas.

Tyrimo tikslas – pritaikyti statistinius modelius ar modelį, kuriais apibrėžtu tikslumu (paklaida <15 proc.) būtų galima prognozuoti pasirinktų atliekų susidarymą.

Tyrimo uždaviniai:

1. Išanalizuoti ir tarpusavyje palyginti mokslinėje literatūroje aprašytus atliekų susidarymo prognostinius metodus, identifikuoti efektyviausius;

2. Išanalizuoti modelių naudojimo mokymo ir modelių testavimo etapų rezultatus bei nustatyti, kurie statistiniai modeliai yra labiausiai tinkami prognozuojant konkrečių atliekų susidarymą;

3. Naudojant pasirinktus modelius parengti medicininių, automobilinių ir dienos šviesos lempų atliekų susidarymo prognozes iki 2020 m.

4. Atlikti nustatytų ir prognozuotų medicininių, automobilinių ir dienos šviesos lempų atliekų susidarymo duomenų jautrumo analizę, palyginti faktinius ir prognostinius rezultatus.

Naujumas: atlikto tyrimo metu buvo palygintos trylikos skirtingų statistinio modeliavimo metodų veikimo galimybės ir efektyvumas prognozuojant medicininių, automobilinių, dienos šviesos lempų ir gyvsidabrio turinčių atliekų kiekių susidarymą. Taip pat buvo išnagrinėtos šių atliekų susidarymo prognozių galimybės naudojant arba nenaudojant atliekų susidarymui galimai įtaką darančius rodiklius. Tai atskleidė, kad, tai pačiai situacijai vertinti naudojant kelis skirtingus modelius, jais gautų prognozių tikslumas gali būti nevienodas. Tokio plataus spektro tyrimas, ieškant panašumų tarp skirtingų ir literatūroje mažai nagrinėtų arba nenagrinėtų atliekų tipų susidarymo tendencijų, buvo atliktas pirmą kartą.

Tikimasi, kad šis mokslinis darbas paskatins atlikti išsamesnius ir detalesnius tyrimus pavojingų ir kitų specifinių atliekų susidarymo prognozavimo srityje ir jų metu bus parengti detalesni ir unikalūs prognozavimo modeliai.

1. LITERATŪROS APŽVALGA

Atlikus prieinamų mokslinės literatūros šaltinių apžvalgą paaiškėjo, kad statistinio modeliavimo metodai dažniausiai buvo naudojami siekiant prognozuoti komunalinių atliekų susidarymą. Pradiniame tyrimo etape pavyko rasti tik kelis straipsnius apie medicininių atliekų susidarymo prognozavimą, o informacija apie kitų pavojingų atliekų susidarymo prognozavimą buvo ypač ribota.

Siekiant prognozuoti atliekų susidarymą, dažniausiai naudojami du pagrindiniai metodai – statistinis matematinis modeliavimas ir sisteminis (dinaminis) modeliavimas. Statistinio modeliavimo metodai buvo pasirinkti dėl galimybės į prognozavimo procesą inkorporuoti socialinius bei ekonominius rodiklius ir pritaikyti automatinio arba kompiuterinio mokymosi (angl. *machine learning*) principus.

Norima pabrėžti, kad šioje disertacijoje nenaudojami ir neanalizuojami metodai, kuriais galima modeliuoti optimalią, kaštų atžvilgiu racionalią ir efektyvią atliekų tvarkymo ar vertinimo sistemą.

Disertacijoje trumpai aptariama ir palyginama 50 prieinamoje mokslinėje literatūroje aprašytų atliekų susidarymo modeliavimo metodų. Tai leido įvertinti šių metodų stiprybes, trūkumus bei galimo panaudojimo aplinkybes.

Disertacijos tyrimo metu naudoti statistinio modeliavimo metodai pateikiami 1.1 lentelėje.

1.1. Statistiniai matematinio modeliavimo metodai

Disertacijoje aprašomų tyrimų metu buvo nuspręsta naudoti šešių skirtingų tipų prognozavimo metodikas, kurioms dažnai buvo taikomas daugiau nei vienas modelis (žr. 1.1 lentelę).

1.1 lentelė. Tyrimuose taikyti matematinio prognozavimo metodai, modeliai ir programinė įranga

Metodas	Modelis	Programinė įranga
Dirbtinių neuronų tinklai (angl. ANN)	Apibendrintos regresijos ir tikimybiniai neuroniniai tinklai (PN/GRN)	Neural Tools 6.

	Daugiasluksnis, tiekiantis į priekį tinklas (MLF)	
Daugialypė tiesinė regresija (angl. MLR)	Enter metodas	SPSS
Dalinių mažiausių kvadratų metodas (angl. PLS)		MiniTab
Atraminių vektorių metodas (angl. SVM)	Klasifikacija	R studio, package e1071
	Nelinijinė regresija	
Neparametrinės regresijos	Bendrieji adityvai	SAS university edition
	Lokali regresija	
	Glodinantys splainiai	
	Branduolinė regresija	
Laiko eilutės (angl. TS)	Slenkantis vidurkis (angl. MA)	StatTools
	Sezoninis eksponentinis glodinimas (angl. SES)	
	Holt eksponentinis glodinimas (angl. Holt's)	

2.TYRIMŲ METODIKA

2.1. Duomenų jautrumo analizė

Anot mokslinės literatūros, nėra vieno geriausio ar pranašiausio metodo analizuojant duomenų jautrumą – kiekvienas metodas turi tam tikrų pranašumų ar trūkumų. Dėl šios priežasties rekomenduojama naudoti bent kelis skirtingo veikimo principo metodus, kurie leistų užtikrinti duomenų jautrumo analizės rezultatų patikimumą. Spearmano ranginė koreliacija (Spearmano rho testas), taškinės diagramos, daugialypė tiesinė regresija (MLR) arba dalinių mažiausių kvadratų metodas (PLS) buvo naudojami atliekant jautrumo analizę šiame tyrime. Atliekant Spearmano rho testą, MRL modeliavimą, buvo naudota SPSS programinė įranga, taškinėms diagramoms – MS Excel programa, PLS – MiniTab programinė įranga. Verta paminėti, kad MLR ir PLS metodų rezultatai turėtų būti traktuojami tik kaip bendrosios gairės, nes daugelyje nagrinėtų duomenų imčių kintamieji koreliavo tarpusavyje, ir tai galėjo iškraipyti gautus rezultatus.

Išsami duomenų jautrumo analizė buvo atlikta tik su tomis parengtomis duomenų imtimis, kurias naudojant modelių mokymo ir testavimo etapai davė teigiamus rezultatus.

2.2. Modelių veikimo tikslumo vertinimas

Pirmas etapas. Siekiant įvertinti atliekų susidarymui prognozuoti naudojamų modelių generuojamų prognozių tikslumą, buvo nuspręsta naudoti keturis vertinimo kriterijus: vidutinė kvadratinė paklaida (RMSE) (žr. 1 lygtį), determinacijos koeficientas (R^2) (žr. 2 lygtį), vidutinė absoliutinė paklaida (MAE) (žr. 3 lygtį) ir vidutinė absoliutinė procentinė paklaida (MAPE) (žr. 4 lygtį).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_0 - Y_p)^2} \quad (1)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_0 - Y_p)^2}{\sum_{i=1}^n (Y_0 - Y_{ave})^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_0 - Y_p| \quad (3)$$

$$MAPE = 100 \text{ proc.} \times \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_0 - Y_p}{Y_0} \right| \quad (4)$$

Čia:

Y_0 – realus susidariusių atliekų kiekis;

Y_{ave} – vidutinis tam tikrų atliekų susidarymo kiekis;

Y_p – prognozuojamas atliekų kiekis.

R^2 reprezentuoja modelio paaiškintą bendrą dispersijos proporciją. MAE yra labiausiai absoliutus ir santykinis reikšmingas modelio klaidos matas, o RMSE – priemonė likusiai matavimo dispersijai, kuri nėra paaiškinta modeliu. MAPE nustato prognozavimo metodo prognozių tikslumą, kuris paprastai išreiškiamas procentais.

Antras, trečias ir ketvirtas etapai. Siekiant įvertinti statistinių modeliavimo metodų efektyvumą antrame tyrimų etape, prognozių tikslumui įvertinti buvo naudojama vidutinė absoliuti procentinė paklaida (MAPE) (žr. 4 lygtį). Dėl antrame tyrimų etape naudotų duomenų imčių specifinių charakteristikų buvo nuspręsta atsisakyti trijų anksčiau naudotų matematinių indeksų.

Nustatyta didžiausia leistina MAPE kritinė vertė yra 15 proc. Modeliai, kurių prognozių paklaidos yra didesnės nei 15 proc., yra pašalinami iš tolesnių tyrimų.

Prognozių absoliutinės paklaidos trečiame tyrimų etape buvo naudojamos vertinant galimą modelių stabilumą ir patikimumą. Šiame etape buvo

apskaičiuotos medicininių, automobilių ir dienos šviesos lempų atliekų susidarymo prognozės iki 2020 metų. Laiko eilučių modeliavimo metodai ilgesniu nei 5 metų laikotarpiu suvienodina generuojamas prognozes, ir jos tampa paskutinių nustatytų ar prognozuotų verčių vidurkiais, todėl buvo nuspręsta apskaičiuoti tik trumpalaikes prognozes (nuo 2016 m iki 2020 m. imtinai).

3. DUOMENŲ IMČIŲ KŪRIMAS

Šiame tyrime naudojamų duomenų imtims kurti buvo naudojami laisvai prieinami Lietuvos statistikos departamento, Aplinkos apsaugos agentūros, Vyriausybės institucijos „Regitra“ ir Europos Komisijos Energetikos ir transporto generalinio direktorato 2004–2015 m. duomenys. Laisvai prieinami oficialūs duomenys buvo naudojami siekiant užtikrinti, kad parengtais prognozavimo metodais ateityje galėtų naudotis tuo suinteresuoti privataus ir viešojo sektorių atstovai.

Esama pavojingų atliekų surinkimo ir apskaitos sistema Lietuvoje buvo sukurta tik 2004 m., todėl ji suteikia galimybę gauti ribotą skaičių aktualių atliekų susidarymo metinių stebėjimų. Nacionaline įstatymų baze reglamentuojama, kad šiai apskaitos sistemai būtų pateikiami tik metiniai pavojingų atliekų surinkimo duomenys, todėl nebuvo įmanoma gauti atitinkamų nacionalinių duomenų apie savaitinį, mėnesinį ar ketvirtinį atliekų susidarymą.

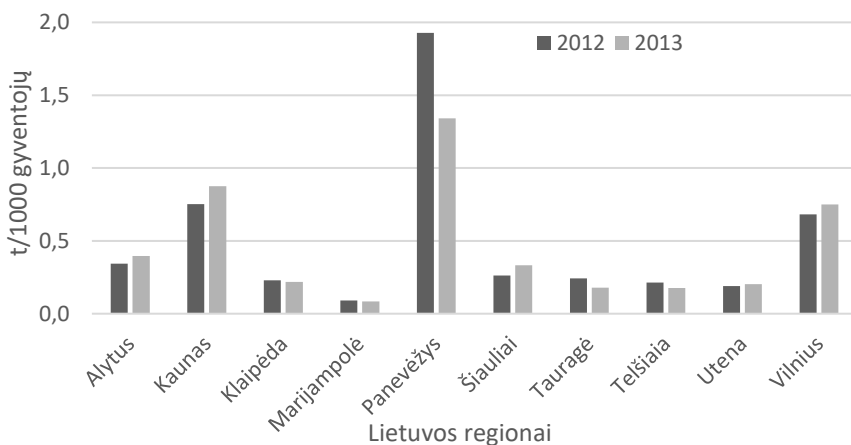
Bandymai surinkti tuos pačius duomenis apie kitas Europos Sąjungos šalis buvo atlikti siekiant išbandyti parengtų modelių pritaikymo galimybes už Lietuvos ribų. Deja, nebuvo rasta viešai prieinamų ir oficialių statistinių duomenų apie dominančių atliekų susidarymą tokiose šalyse, kaip Didžioji Britanija, Danija, Vokietija, Latvija, Graikija, Estija. Minėtos šalys daugiausia skelbia duomenis apie komunalinių ar bendrųjų atliekų susidarymą – galimybės surinkti išsamesnius duomenis apie medicininių, dienos šviesos lempų ar automobilių atliekų susidarymą yra ribotos.

3.1. Medicininių atliekų duomenų imtis

Surinkus ir išanalizavus visus galimus medicininių atliekų duomenis, buvo nuspręsta parengti tris duomenų imtis: 1) 2012–2013 m. metiniai regioniniai duomenys; 2) 2004–2013 m. nacionaliniai metiniai suminiai medicininių atliekų susidarymo duomenys; 3) 2004–2013 m. nacionaliniai metiniai pavojingų medicininių atliekų susidarymo duomenys.

Regionų duomenų imtį sudaro metiniai duomenys iš 10 Lietuvos regionų 2012–2013 m.: Alytaus, Kauno, Klaipėdos, Marijampolės, Panevėžio, Šiaulių,

Tauragės, Telšių, Utenos ir Vilniaus. Iš viso 20 atvejų su 24 nepriklausomais ir priklausomais kintamaisiais. Šie duomenys buvo normalizuoti ir perskaičiuoti 1000 gyventojų. Spearmano racionalumo koreliacijos koeficientai apskaičiuoti siekiant užtikrinti, kad medicininėms atliekoms prognozuoti būtų naudojami tik labiausiai jų susidarymui įtaką darantys rodikliai. Koeficientai parodė, kad 5 kintamieji reikšmingai koreliuoja su medicininių atliekų susidarymu – apsilankymai ligoninėse ir kitose medicinos įstaigose, vaikų skaičius regione, lovų skaičius ligoninėse ir kitose medicinos įstaigose, vidutinė gyvenimo trukmė ir gydytojo vizitų skaičius tame regione. Šie kintamieji buvo naudojami kaip nepriklausomi kintamieji atliekant tolesnius skaičiavimus pasirinkta programine įranga.



3.1.1 pav. Duomenys apie medicininių atliekų susidarymą Lietuvos regionuose 2012–2013 m., tonos/1000 gyventojų (Aplinkos apsaugos agentūra, 2015)

Dideli statistinių duomenų skirtumai tarp regionų su didžiausiais ir mažesniais miestais susiformavo atlikus duomenų perskaičiavimą 1000 gyventojų, ir tai nulėmė įregistruotų gyventojų skaičius atitinkamuose regionuose.

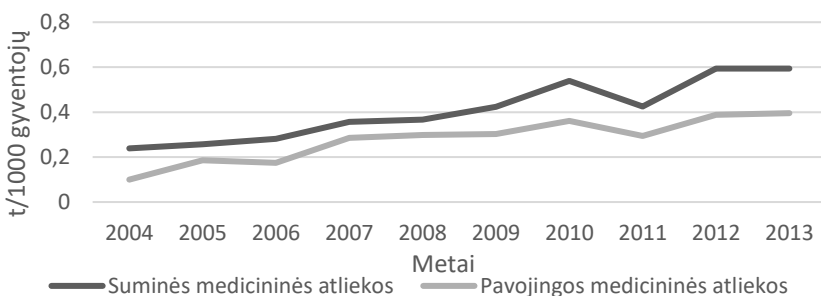
Suminių medicininių atliekų duomenų imtis yra sudaryta iš 10 atvejų, kurių kiekviename yra 17 kintamųjų. Suminių medicininių atliekų susidarymas yra priklausomas kintamasis. Šie duomenys buvo normalizuoti ir perskaičiuoti 1000 gyventojų. Tada buvo apskaičiuoti Spearmano racionalumo koreliacijos koeficientai, siekiant užtikrinti, kad suminėms medicininėms atliekoms prognozuoti būtų naudojami tik didžiausią įtaką jų susidarymui darantys rodikliai. Šis veiksmas parodė, kad efektyviausi nepriklausomi kintamieji galėtų būti: vizitai

klinikose, naujagimių skaičius, senyvų žmonių skaičius, vaikų skaičius, lovų skaičius ir lovų apyvarta ligoninėse, vidutinė gyvenimo trukmė, gydytojų skaičius, pacientų skaičius, lojadienių skaičius, dieninių atvejų skaičius sveikatos priežiūros įstaigose, vidutinis ligoninėje ir stacionare praleistas ligonio laikas, AIDS, ŽIV ir vėžio atvejai.

Visų medicininių atliekų susidarymo kiekiai pateikti 3.1.2 pav.

Pavojingų medicininių atliekų duomenų imtyje priklausomas kintamasis yra pavojingų medicininių atliekų susidarymas. Rinkinys susideda iš 10 stebėjimų su 16 išvardytų nepriklausomų kintamųjų.

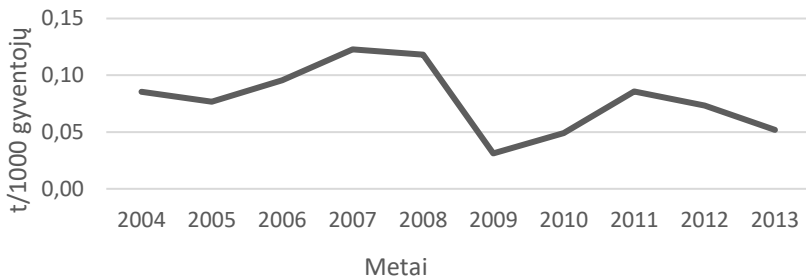
Susidarantys suminių ir pavojingų medicininių atliekų kiekiai pateikti 3.1.2 pav.



3.1.2 pav. Suminių ir pavojingų medicininių atliekų susidarymas 2004–2013 m., t/1000 gyventojų (Aplinkos apsaugos agentūra, 2015)

3.2. Dienos šviesos lempų ir kitų gyvsidabrio turinčių atliekų duomenų imtis

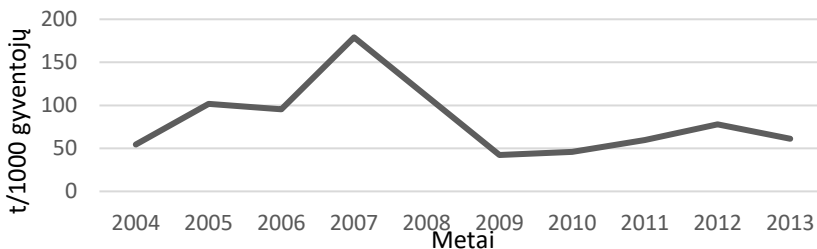
Du sudaryti duomenų rinkiniai susideda iš 10 stebėjimų, kurie skiriasi tik nepriklausomų kintamųjų skaičiumi: vienas papildomai apima kintamuosius, susijusius su namų ūkio vartojimu (iš viso 7 nepriklausomi kintamieji), o kitas – ne (iš viso 4 nepriklausomi kintamieji). Kintamieji, bendri abiem duomenų rinkiniams, kurie galimai daro įtaką dienos šviesos lempų susidarymui ir atliekoms, kurių sudėtyje yra gyvsidabrio: elektros energijos galia, išlaidos elektros energijai, BVP ir pramonės metinė apyvarta. Kintamieji, susiję su namų ūkių išlaidomis: išlaidos, susijusios su elektros, dujų ir kito kuro naudojimu, išlaidos, susijusios su įprastine namų ūkių priežiūra, visos namų ūkių išlaidos ekonominėje teritorijoje ir užsienyje. Visi kintamieji buvo perskaičiuoti taip, kad atitiktų 1000 gyventojų. Spearmano ir Pearsono koreliacijos skaičiavimas parodė, kad nė vienas iš naudotų nepriklausomų kintamųjų nėra labai susijęs su analizuotų atliekų susidarymu.



3.2.1 pav. Dienos šviesos lempų ir kitų gyvsidabrio turinčių atliekų susidarymas 2004–2013 m., t/1000 gyventojų (Aplinkos apsaugos agentūra, 2015)

3.3. Automobilinių atliekų duomenų imtis

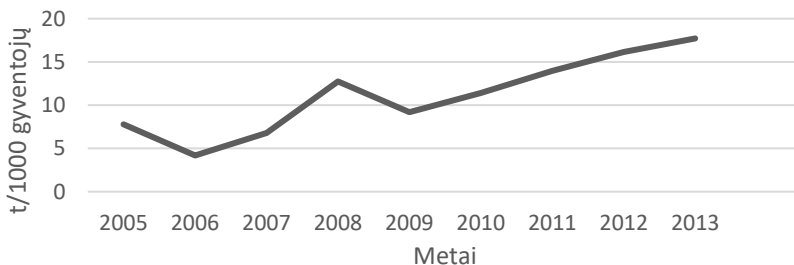
Buvo sudaryti du automobilinių atliekų duomenų rinkiniai. Bendrasis metinis automobilių atliekų susidarymas yra priklausomas kintamasis, ir 5 socialiniai bei ekonominiai rodikliai laikomi priklausomais kintamaisiais suminiame automobilių atliekų duomenų rinkinyje. Šie rodikliai buvo atrinkti iš didesnio pirminio duomenų rinkinio skaičiuojant Spearmano ranginę koreliaciją, ir ji parodė, kurie rodikliai gali turėti didžiausią įtaką automobilių atliekų susidarymui. Nepriklausomi kintamieji suminių automobilių atliekų duomenų rinkinyje yra: bendras naujų automobilių skaičius, bendras naujų automobilių skaičius (M1 tipas), keleivių, kurie važiuoja automobiliais, nukeliautas atstumas, keleivių, važiuojančių autobusais, nukeliautas atstumas. Visi kintamieji buvo perskaičiuoti taip, kad atitiktų 1000 gyventojų. Suminių duomenų rinkinį sudaro 6 kintamieji ir 10 stebėjimų.



3.3.1 pav. Nacionaliniai suminių automobilinių atliekų susidarymo duomenys 2004–2013 m., t/1000 gyventojų (Aplinkos apsaugos agentūra, 2015)

Ryškus automobilių atliekų susidarymo sumažėjimas 2009 m. siejamas su ekonominės krizės pradžia 2008 metais. Ekonominiai rodikliai, kurie galimai daro įtaką automobilių atliekų susidarymui (BVP, registruotų automobilių skaičius, kelionių automobiliais ir autobusais skaičius), 2009 m. taip pat smarkiai sumažėjo, todėl manoma, kad šie duomenų svyravimai tinkamai įvertinami mokant modelius su parengta duomenų imtimi. Tai pagrindžia disertacijos 3 priede, 6 ir 7 lentelėse, pateikiami ekonominiai-socialiniai indikatoriai.

Antroje duomenų imtyje metinis pavojingų automobilių atliekų susidarymas buvo priklausomas kintamasis. Šioje duomenų imtyje naudojama 13 socialinių ir ekonominių rodiklių, kurie traktuojami kaip nepriklausomi kintamieji. Šie duomenys buvo normalizuoti ir perskaičiuoti 1000 gyventojų. Tada buvo apskaičiuoti Spearmano racionalumo koreliacijos koeficientai, siekiant užtikrinti, kad pavojingų automobilių atliekų kiekiui prognozuoti būtų naudojami tik didžiausią įtaką atliekų susidarymui turintys rodikliai. Nepriklausomi kintamieji pavojingų automobilių atliekų duomenų rinkinyje yra: įregistruotų naudotų automobilių skaičius (M1 tipas), išregistruotų transporto priemonių skaičius, bendras įregistruotų automobilių skaičius (M1 tipas), įregistruotų lengvųjų automobilių skaičius (M1-M3 tipai), įregistruotų mopedų skaičius (L1-L2 tipai), įregistruotų priekabų skaičius (O1-O4 tipai), bendras įregistruotų transporto priemonių skaičius, eismo įvykių skaičius, šalies BVP, tarptautiniai ir bendri pervežimai transporto priemonėmis, įregistruotomis Lietuvoje. Pavojingų automobilių atliekų duomenų imtį sudaro 14 kintamųjų ir 10 stebėjimų.



3.3.2 pav. Nacionaliniai pavojingų automobilių atliekų susidarymo duomenys 2004–2013 m., t/1000 gyventojų (Aplinkos apsaugos agentūra, 2015)

3.4. Laiko eilučių bandymai trečiame etape

Trečiame tyrimo etape iš 2004–2015 m. statistinių atliekų susidarymo duomenų buvo sudaryta 10 skirtingo ilgio duomenų imčių, kurias naudojant buvo

išbandytas laiko eilučių Holt modelių prognozavimo tikslumas, įvedant skirtingo ilgio duomenų imtis. Imčių sudarymo metodika pateikta 3.4.1 lentelėje.

3.4.1 lentelė. Holt metodu išbandytų skirtingo ilgio duomenų imčių kūrimo principas

Duomenų imtis	Stebėjimų skaičius	Periodas	Prognozių skaičius iki 2015 m.
Nr.1	1	2004	11
Nr. 2	2	2004–2005	10
Nr. 3	3	2004–2006	9
Nr. 4	4	2004–2007	8
Nr. 5	5	2004–2008	7
Nr. 6	6	2004–2009	6
Nr. 7	7	2004–2010	5
Nr. 8	8	2004–2011	4
Nr. 9	9	2004–2012	3
Nr. 10	10	2004–2013	2

Išbandant skirtingo ilgio duomenų imtis TS MA metodu paaiškėjo, kad tik 10 stebėjimų duomenų imtis gali teikti teigiamus rezultatus, todėl trečiame tyrimo etape buvo išbandomi tik skirtingos apimties (trijų, keturių ir penkių stebėjimų) laiko eilučių vidurkiai.

4. NAGRINĖJAMŲ ATLIEKŲ SUSIDARYMO IR PRIKLAUSOMYBĖS ANALIZĖ

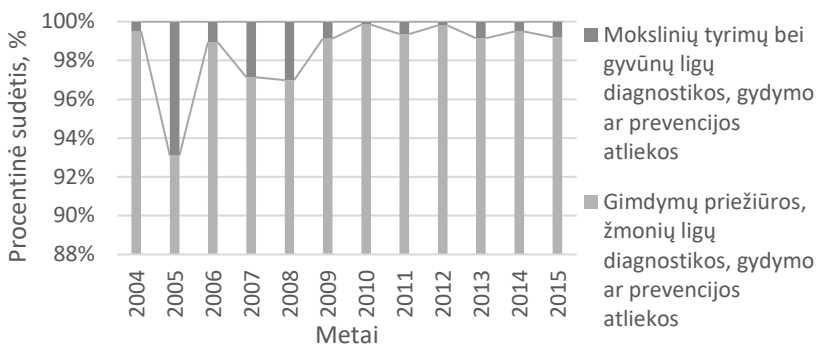
Šioje skiltyje nagrinėjami atliekų susidarymo duomenys nuo 2004 m. iki 2015 m.

Suminės medicininės atliekos

Per nagrinėjamą periodą medicininės atliekos vidutiniškai sudarė apie 0,03 proc. bendro šalyje susidariusio atliekų kiekio tais metais. Skaičiuojama, kad vidutiniškai šių atliekų susidarymas augo 13,5 proc. per metus.

Nagrinėjamu laikotarpiu vidutiniškai 98,7 proc. medicininių atliekų buvo gimdymų priežiūros, žmonių ligų diagnostikos, gydymo ar prevencijos atliekos, vidutinis augimas – 13,56 proc. Apie 61 proc. šios kategorijos atliekų sudarė atliekos, kurių rinkimui ir šalinimui taikomi specialūs reikalavimai, siekiant išvengti infekcijos, 20,68 proc. – atliekos, kurių rinkimui ir šalinimui netaikomi

specialūs reikalavimai siekiant išvengti infekcijos, 8 proc. – vaistai, kuriems nepriskiriamas 18 01 08 kodas.



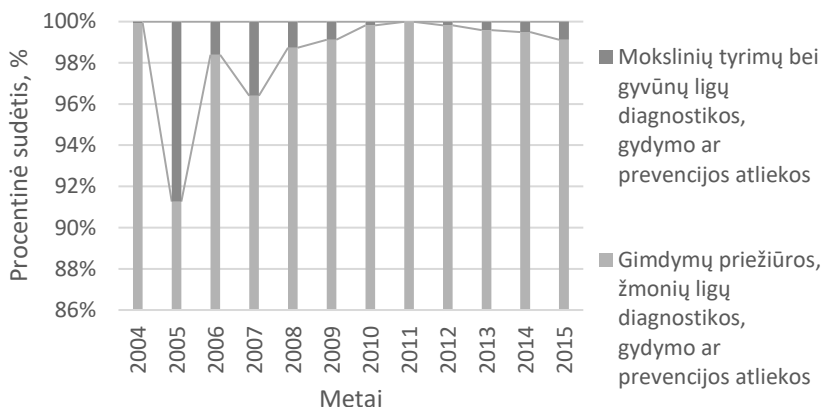
4.1 pav. Medicininių atliekų sudėties pokyčiai 2004–2015 m.

Mokslinių tyrimų bei gyvūnų ligų diagnostikos, gydymo ar prevencijos atliekos vidutiniškai sudarė 1,53 proc. medicininių atliekų kasmet. Nagrinėjamu periodu buvo fiksuojami dideli šių atliekų kiekių svyravimai – nuo 0,16 proc. iki 6,90 proc. Cheminės medžiagos, sudarytos iš pavojingų cheminių medžiagų arba jų turinčios, buvo bene vienintelės atliekos, kurių kasmet būdavo surenkamas maždaug vienodas kiekis. Kitų tipų atliekų susidarymas nebuvo vienodas, ir galėtų būti siejamas su vykdomais didelio masto tyrimais ar naminių bei laukinių gyvūnų ligomis.

Taip pat buvo pastebėta, kad 2004–2005 m. pasikeitė medicininių atliekų klasifikacija – smarkiai sumažėjo atliekų, kurioms netaikomi specialūs reikalavimai dėl galimų infekcijų, ir citotoksinių ar citostatinių vaistų atliekų kiekiai (atitinkamai nuo 58,72 proc. iki 13,57 proc. ir nuo 40,61 proc. iki 5,38 proc., vėliau neviršijo 1 proc. ribos). Iškart po to 2006 m. išaugo medicininių atliekų, kurių rinkimui ir šalinimui taikomi specialūs reikalavimai, kad būtų išvengta infekcijos (nuo 0 proc. iki 60,35 proc.), ir cheminių medžiagų, nenurodytų 18 01 06 (nuo 0,08 proc. iki 9,44 proc., vėlesniais metais sudarė vidutiniškai 8 proc.), susidarymas. Šie pokyčiai siejami su Lietuvos įstojimu į Europos Sąjungą, kuri įpareigojo perkelti atliekų tvarkymo normatyvus į nacionalinę teisę ir praktiką.

Pavojingos medicininės atliekos

Nagrinėjamu periodu vidutinis metinis pavojingų medicininių atliekų susidarymas augo 19,19 proc. ir vidutiniškai sudarė 0,53 proc. šalyje susidarančių pavojingų atliekų. Vidutiniškai 98,47 proc. pavojingų medicininių atliekų susidarė gimdymų priežiūros, žmonių ligų diagnostikos, gydymo ar prevencijos sektoriuje. Atliekos, kurių rinkimui ir šalinimui taikomi specialūs reikalavimai siekiant išvengti infekcijos, sudarė vidutiniškai 96 proc. visų pavojingų medicininių atliekų (vidutinis metinis augimas – 9,32 proc.). 1,2 proc. sudarė cheminės medžiagos, sudarytos iš pavojingų cheminių medžiagų arba jų turinčios, ir 0,94 proc. citotoksinių ar citostatinių vaistų.



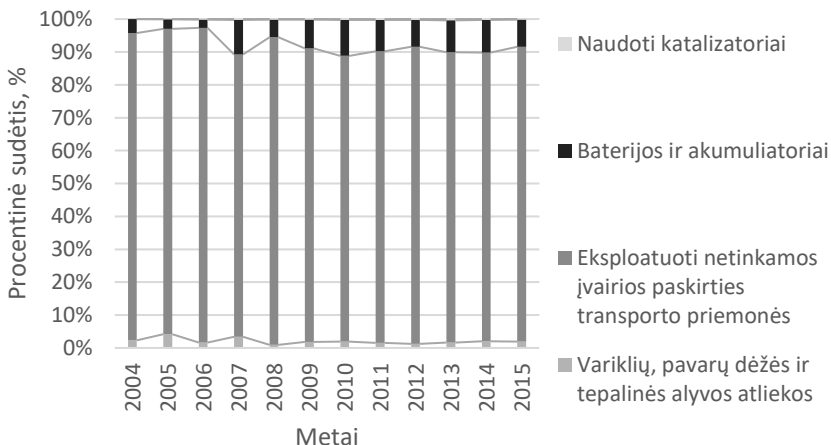
4.2 pav. Pavojingų medicininių atliekų sudėties pokyčiai 2004–2015 m.

Mokslinių tyrimų bei gyvūnų ligų diagnostikos, gydymo ar prevencijos atliekos vidutiniškai sudarė 1,53 proc. bendro pavojingų medicininių atliekų metinio fiksuojamo kiekio. 2005 m. matomas didelis šių atliekų kiekio išaugimas, kuris yra siejamas su atliekomis, kurių rinkimui ir šalinimui taikomi specialūs reikalavimai, kad būtų išvengta infekcijos. Tokie nepastovūs svyravimai galėtų būti siejami su vykdomais didelio masto tyrimais ar naminių bei laukinių gyvūnų ligomis. Taip pat pastebima, kad tik nuo 2009 m. medicininės atliekos iš mokslinių tyrimų su gyvūnais pradėtos skirstyti į atskiras smulkesnes kategorijas.

Suminės automobilinės atliekos

2004–2015 m. automobilinės atliekos sudarė vidutiniškai 3,93 proc. visų metinių atliekų, vidutinis metinis augimas buvo 9,41 proc. Į šį tyrimą įtrauktos keturios pagrindinės automobilių atliekų kategorijos: variklių, pavarų dėžės ir

tepalinės alyvos atliekos (7,76 proc. metinis augimas), eksploatuoti netinkamos įvairios paskirties transporto priemonės (8,99 proc. metinis augimas), baterijos ir akumuliatoriai (16,21 proc. metinis augimas) ir panaudoti katalizatoriai (105,44 proc. metinis augimas). Pastarųjų atliekų dalis bendroje automobilių atliekų sumoje yra labai nedidelė, palyginti su susidaranciais kitų kategorijų atliekų kiekiais (4.3 pav.)

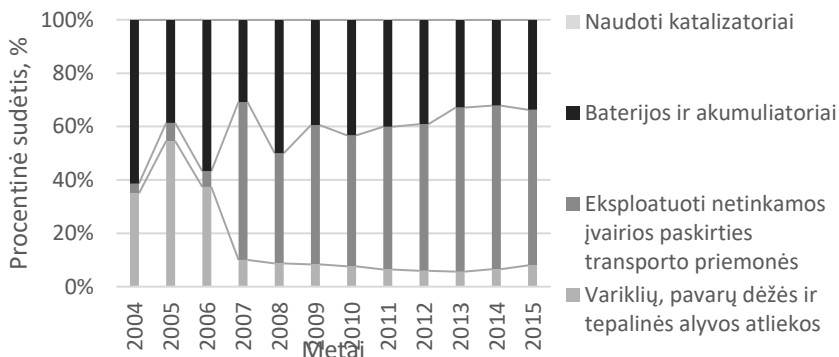


4.3 pav. Automobilių atliekų sudėties pokyčiai 2004–2015 m.

Variklių, pavarų dėžės ir tepalinės alyvos atliekos sudaro vidutiniškai 2,04 proc. metinių automobilių atliekų kiekio, eksploatuoti netinkamos įvairios paskirties transporto priemonės – 90 proc. (71,54 proc. minimo kiekio yra magnetiniai metalai), baterijos ir akumuliatoriai – 7,57 proc. Nuo 2006 m. pastebimas panaudotų katalizatorių atliekų augimas siejamas su automobilių standartų pasikeitimais, jaunėjančiu nacionaliniu automobilių parku ir išpopuliarėjusiomis katalizatorių perdirbimo technologijomis, tad dabar šios kategorijos atliekos kasmet vidutiniškai sudarytų apie 0,2 proc. visų automobilių atliekų.

Pavojingos automobalinės atliekos

Nuo 2007 m. vidutiniškai 7,48 proc. pavojingų automobilių atliekų sudaro variklių, pavarų dėžės ir tepalinės alyvos atliekos, pastebimas kasmetinis mažėjimas yra 2,71 proc. Šie pokyčiai nuteikia optimistiškai, nes alyvų ir tepalų perdirbimas yra technologiškai sunkus ir brangus procesas.



4.4 pav. Pavojingų automobilių atliekų sudėties pokyčiai 2004–2015 m.

Nuo 2007 m. įvykusių pokyčių automobilių atliekų apskaitos sistemoje vidutiniškai 54,57 proc. pavojingų automobilių atliekų sudaro eksploatuoti netinkamos įvairios paskirties transporto priemonės, kurių 51 proc. sudaro eksploatuoti netinkamos transporto priemonės. Vidutiniškai 41,52 proc. pavojingų automobilių atliekų sudaro baterijų ir akumuliatorių atliekos ir, nepaisant 16,8 proc. kasmetinio šių atliekų prieaugio, jų dalis bendrame kontekste išlieka maždaug tolygi.

Dienos šviesos lempos ir kitos gyvsidabrio turinčio atliekos

Vidutiniškai šios kategorijos atliekos kasmet augo po 3,5 proc. ir sudarė apie 0,11 proc. bendro atliekų kiekio tais metais. 2004–2015 m. šių atliekų dalis bendrame atliekų sraute kito 0,07–0,19 proc.

4.1. Duomenų jautrumo analizė

Regioninės medicininės atliekos

Dvidešimt įvairių socialinių-ekonominių rodiklių buvo išbandyti naudojant Spearmano rho testą, tačiau regioninių medicininė atliekų atveju tik 5 kintamieji teigiamai koreliavo su nagrinėjamomis atliekomis. Šie kintamieji buvo naudoti sudarant regioninių medicininė atliekų duomenų imtį. Spearmano rho testas atskleidė, kad vizitai į ligonines, lovų skaičius ligoninėse, vidutinė gyvenimo trukmė ir gydytojų vizitai, palyginti su minimų atliekų rodikliu, rodo teigiamą koreliaciją. Rajone esančių vaikų skaičius parodė neigiamą koreliaciją su regioninėmis medicininėmis atliekomis. Tai leidžia daryti prielaidą, kad jaunėjanti gyventojų bendruomenė sudaro sąlygas medicininė atliekų susidarymo mažėjimui tam tikruose regionuose.

Suminės medicininės atliekos

Penkiolika iš dvidešimties išbandytų socialinių-ekonominių rodiklių ryškiai koreliavo su suminėmis medicininėmis atliekomis. Keturi iš penkiolikos kintamųjų neigiamai koreliavo. Svarbu paminėti, kad visi pasižymėję kintamieji taip pat ryškiai koreliavo tarpusavyje.

Taškinės diagramos šiam atvejui nebuvo paruoštos, nes, naudojant TS Holt metodą, gauti geriausi rezultatai, ir šio metodo veikimui nereikia pateikti socialinių-ekonominių kintamųjų.

Pavojingos medicininės atliekos

Sudarant šią duomenų imtį buvo naudojami tie patys nepriklausomi kintamieji, kaip ir suminių medicininių atliekų atveju, ir Spearmano rho testo rezultatai beveik nesiskyrė. Visi pasižymėję kintamieji vėl ryškiai koreliavo tarpusavyje.

4.1.1 lentelė. Pavojingų medicininių atliekų susidarymo duomenų jautrumo rezultatai, remiantis CC ir taškinių diagramų skaičiavimais

Kintamasis	CC		SPR ²	
	Nustatyta vertė	SVM regresijos vertė	Nustatyta vertė	SVM regresijos vertė
Vizitai klinikose	0,87**	1,00**	0,81	0,78
Naujagimiai	0,85**	1,00**	0,84	0,57
Vaikai	-0,91**	-1,00**	0,71	0,35
Pensininkai	0,89**	0,70	0,89	0,35
Lovų apyvarta ligoninėse	0,92**	0,15	0,32	0,00
Lovų skaičius ligoninėse	0,91**	0,30	0,69	0,12
Vidutinė gyvenimo trukmė	0,75*	0,90*	0,69	0,08
Gydytojų vizitai	0,86**	1,00**	0,63	0,59
Pacientų ligoninėse skaičius	0,69*	0,90*	0,75	0,86
Pacientų nakvynių ligoninėse skaičius	-0,95**	-1,00**	0,63	0,79

Dienos atvejų ligininėse skaičius	0,91**	0,50	0,9	0,54
Vidutinė stacionare praleista trukmė	-0,86**	-0,97**	0,86	0,35
Vidutinė ligininėje praleista trukmė	-0,94**	-0,73	0,82	0,18
AIDS atvejų skaičius	-0,94**	1,00**	0,74	0,64
ŽIV atvejų skaičius	0,91**	1,00**	0,80	0,63
Vėžio atvejų skaičius	0,91**	–	0,87	0,83

* 0,05 patikimumo / reikšmingumo lygmuo

** 0,01 patikimumo / reikšmingumo lygmuo

Reikšmingi rezultatai pažymėti pilkais.

Remiantis koreliacijos koeficientų skaičiavimais, aštuoni socialiniai-ekonominiai indikatoriai tapo labai svarbūs SVM modelio sudarymui ir gebėjimui generuoti taiklias prognozes. Penki naudoti kintamieji nebuvo tokie aktualūs SVM modelio sudarymui. Taškinių diagramų analizė parodė, kad trys nepriklausomi kintamieji naudojami kaip SVM modelio pagrindas prognozuojant pavojingas medicininės atliekas. Du iš trijų kintamųjų sutampa su tais, kurie buvo atrinkti CC skaičiavimų metu.

Fiksuojamas didelis duomenų jautrumas SVM modelio veikimui nėra pageidautinas faktorius, ir tai gali neigiamai paveikti pavojingų medicininių atliekų prognozių tikslumą. Kadangi dėl modelio specifikacijos tik 7 iš 10 stebėjimų buvo panaudoti modelio mokymui (nors ir su tokia trumpa imtimi gauti teigiami rezultatai), tokios pačios struktūros modelis galėtų būti iš naujo mokomas naudojant jau visus turimus duomenis. Tikėtina, kad tai sumažintų SVM modelio jautrumą imties duomenims.

Suminės automobalinės atliekos

Tik keturi iš dvidešimt išbandytų nepriklausomų kintamųjų aiškiai koreliavo su suminėmis automobalinėmis atliekomis ir įveikė Spearmano rho testą. Rezultatai parodė, kad tarpusavyje nekoreliuoja tik skirtingomis transporto priemonėmis keliautojų nukeliautus atstumus pristatantys kintamieji.

4.1.2 lentelė. Suminių automobilių atliekų susidarymo duomenų jautrumo rezultatai, remiantis CC, taškinių diagramų ir regresijos metodų skaičiavimais

Kintamasis	CC		SPR ²		Regresija	
	Nustatyta vertė	Branduolinės regresijos vertė	Nustatyta vertė	Branduolinės regresijos vertė	MLR	PLS
Naujos transporto priemonės	0,80**	0,67*	0,52	0,56	3,49	3,49
Nauji automobiliai	0,82**	0,67*	0,49	0,56	– 2,88	2,89
Keleivių nukeliautas atstumas automobiliais	0,66**	0,62*	0,24	0,34	– 0,04	0,04
Keleivių nukeliautas atstumas autobusais	0,82*	0,75**	0,54	0,57	0,44	0,44

* 0,05 patikimumo / reikšmingumo lygmuo

** 0,01 patikimumo / reikšmingumo lygmuo

Skaičiuojant CC koeficientus branduolinės regresijos prognozėms, paaiškėjo, kad nepriklausomų kintamųjų įtaka sumažėjo 0,1–0,2 punkto, bet išlaikė stiprius koreliacinius ryšius su automobalinėmis atliekomis. Nepriklausomų kintamųjų vertės šiek tiek išaugo, kai jautrumo analizė buvo atlikta naudojant taškines diagramas.

Jautrumo analizė naudojant regresijos metodus parodė, kad didžiausią įtaką modelių veikimui galimai turi naujų transporto priemonių ir naujų automobilių kintamieji. Keleivių nukeliautų atstumų rodikliai nebuvo svarbūs MLR modelio veikimui. MLR jautrumo analizės rezultatai gali būti naudojami tik kaip bendrosios gairės tiriant duomenų jautrumą nagrinėjamų atliekų atveju, nes nepriklausomi kintamieji smarkiai koreliuoja tarpusavyje.

Pavojingos automobulinės atliekos

Trylika iš dvidešimties kintamųjų smarkiai koreliavo su pavojingų atliekų rodikliu po Spearmano rho testo. Testo rezultatai taip pat parodė, kad M1 tipo naujai registruotų dėvėtų automobilių, nurašytų transporto priemonių, iš viso šalyje įregistruotų M1 tipo automobilių skaičius ir eismo įvykių skaičius rodo neigiamą koreliaciją.

Daugelis nepriklausomų kintamųjų pavojingų automobilinių atliekų duomenų imtyje koreliuoja tarpusavyje.

4.1.3 lentelė. Pavojingų automobilinių atliekų susidarymo duomenų jautrumo rezultatai, remiantis CC, taškinių diagramų ir regresijos metodų skaičiavimais

Kintamasis	CC		SPR ²		Regresija
	Nustatyta vertė	Branduolinės regresijos vertė	Nustatyta vertė	Branduolinės regresijos vertė	PLS
Naudoti automobiliai	0,67*	0,54	0,32	0,32	0,85
Nurašytos transporto priemonės	-0,67*	-0,68*	0,51	0,51	0,13
Nurašyti automobiliai	-0,66*	-0,67*	0,50	0,51	1,28
M1	0,92*	0,54	0,15	0,25	2,01
M1–M3	0,92**	0,54	0,14	0,24	1,88
N1–N3	0,95**	0,39	0,03	0,09	-5,59
L1–L7	0,92**	0,73**	0,35	0,46	4,55
O1–O4	0,92**	0,79**	0,54	0,63	-2,04
Visos transporto priemonės	0,92**	0,54	0,17	0,27	1,10
Eismo įvykių skaičius	-0,72*	-0,75**	0,72	0,68	1,05
BVP	0,93**	0,89**	0,89	0,79	3,36
Tarptautiniai pervežimai	0,90**	0,84**	0,84	0,71	-1,93

Tarptautiniai ir nacionaliniai pervežimai	0,90**	0,84**	0,84	0,71	-0,95
---	--------	--------	------	------	-------

* 0,05 patikimumo / reikšmingumo lygmuo

** 0,01 patikimumo / reikšmingumo lygmuo

Reikšmingi rezultatai pažymėti pilkais.

Pakartotinai atlikus CC skaičiavimus ir lyginant su branduolinės regresijos rezultatais paaiškėjo, kad penki kintamieji tapo mažai svarbūs modelio sudarymui (naujai įregistruoti naudoti automobiliai ir M1, M1–M3, N1–N3, bendras transporto priemonių skaičius). Šie kintamieji apima pagrindines transporto priemonių rūšis, kurios taip pat smarkiai koreliuoja tarpusavyje. Kiti svarbūs kintamieji išlaikė savo pozicijas, nors ryškiai teigiamai koreliuojančių kintamųjų įtaka šiek tiek sumenko, o kintamieji su stipria neigiama koreliacija ją dar labiau sustiprino.

Nagrinėjant taškinių diagramų rezultatus buvo pastebėta, kad penkių kintamųjų įtaka paaugo (visi kintamieji susiję su transporto priemonių skaičiumi), trijų kintamųjų koreliacija sumenko 0,1 punkto (visi kintamieji susiję su ekonomine situacija ir stabilumu) ir trijų kintamųjų pozicija nepasikeitė.

PLS rezultatai parodė, kad modelis daugiausiai rėmėsi N1–N3, L1–L7, BVP rodikliais. Kiti didelę įtaką turėję rodikliai buvo naudotų automobilių skaičius, M1 ir O1–O4 rodikliai. Mažiausius svertinius koeficientus turėjo tarptautinių ir nacionalinių pervežimų, nurašytų automobilių ir naudotų mašinų rodikliai.

Dienos šviesos lempos ir kitos gyvsidabrio turinčios atliekos

Dienos šviesos lempų atveju nė vienas iš naudotų socialinių-ekonominių indikatorių stipriai nekoreliavo su nagrinėjamomis atliekomis. Neturint kito pasirinkimo, į tolesnius tyrimus buvo įtraukti visi septyni išbandyti kintamieji. Buvo pastebėta, kad kai kurie nepriklausomi kintamieji smarkiai koreliuoja tarpusavyje. Visos nagrinėtos koreliacijos buvo teigiamos, ir tai parodė, kad kintamieji išlaikė tas pačias augimo tendencijas, greičiausiai būtent šis bruožas ir leido sudaryti duomenų imčiai pasiekti teigiamą rezultatą.

4.1.4 lentelė Dienos šviesos lempų ir kitų gyvsidabrio turinčių atliekų susidarymo duomenų jautrumo rezultatai, remiantis CC, taškinių diagramų ir regresijos metodų skaičiavimais

Kintamasis	CC		SPR ²		Regresija
	Nustatyta vertė	Branduolinės regresijos vertė	MLR	Branduolinės regresijos vertė	PLS
Elektros galia	-0,43	-0,61*	0,24	0,31	-1,05
BVP	-0,13	-0,40	0,08	0,11	0,91
Pramonės metinė apyvarta	-0,13	-0,33	0,06	0,09	2,04
Namų ūkių išlaidos elektrai ir kitam kurui	-0,41	-0,60*	0,29	0,34	-1,26
Namų ūkių išlaidos kasdienei priežiūrai	-0,26	-0,49	0,18	0,22	-0,94
Visos namų ūkių išlaidos	-0,23	-0,47	0,09	0,12	0,12

* 0,05 patikimumo / reikšmingumo lygmuo

** 0,01 patikimumo / reikšmingumo lygmuo

Reikšmingi rezultatai pažymėti pilkai.

Buvo pastebėta, kad, atlikus CC skaičiavimus naudojant branduolinės regresijos sugeneruotas reikšmes, nepriklausomų kintamųjų įtaka ir svarba išaugo, kai kuriais atvejais net pasiekė reikšmingos koreliacijos ribas. Deja, taškinės diagramos neatskleidė tokių pačių teigiamų pokyčių, nors rodiklių, kurie pasiekė reikšmingos koreliacijos ribas, įtaka išliko reikšmingiausia. Sudarant PLS modelį, didžiausias svertinis koeficientas buvo skirtas metinei pramonės apyvartai, ir modelis rėmėsi būtent šio rodiklio stabilumu ir įtaka. Didžiausi neigiami svertiniai koeficientai buvo suteikti namų ūkių išlaidoms elektros energijai ir elektros galiai.

Kadangi PLS modelio likutinės reikšmės nėra išsidėsčiusios pagal normalumo taisyklės, tokie jautrumo analizės rezultatai gali būti laikomi bendrosiomis gairėmis.

5.1 ETAPAS: PIRMINIŲ PROGNOZIŲ IR MODELIŲ VEIKIMO ĮVERTINIMAS

Pirmojo tyrimų etapo metu buvo sudaromi skirtingi modeliai iš pasirinktų ir anksčiau minėtų metodikų. Visos iš 2004–2013 m. duomenų sudarytos imtys buvo naudotos modeliams išbandyti ir vėliau sugeneruotoms teorinėms prognozėms palyginti su realiomis nustatytomis vertėmis. 5.1.1 lentelėje pateikiamas gautų rezultatų apibendrinimas.

5.1.1 lentelė. 1 tyrimų etapo metu geriausiai įvertinti modeliai, kurie buvo naudoti tolesniuose tyrimo etapuose

	ANN		Tiesinė regresija		SVM		Neparametrinė regresija			TS			
	PN-GRN	MLF	MLR	PLS	Klasifikacija	Regresija	Bendrieji adityvai	Glodinantys splainai	Lokali regresija	Branduolinė regresija	MA	SES	Holt
Lempos be namų ūkių	X ₁						X ₁			X ₁	X ₂ *	X ₂ *	X ₂ *
Lempos su namų ūkiais			X ₁	X ₁			X ₁			X ₁	X ₂ *	X ₂ *	X ₂ *
Suminės medicininės atliekos						X				X ₁	X ₂ *	X ₂ *	X ₂ *
Pavojin-gos medicini-						X		X ₁		X ₁	X ₂ *	X ₂ *	X ₂ *

nės atliekos												
Suminės automobalinės atliekos			X ₁ *			X ₁ *	X ₁		X ₁ *	X ₂ *	X ₂ *	X ₂ *
Pavojingos automobalinės atliekos			X ₁ *			X ₁ *	X ₁ *	X ₁ *	X ₁ *	X ₂ *	X ₂ *	X ₂ *

* Šie modeliai buvo pridėti 2 tyrimų etape po to, kai 1 etape gerai įvertinti modeliai, kurių MAPE didesnė nei 15 proc. Pridėtų modelių rezultatai 1 etapo metu buvo šiek tiek prastesni.

X₁ – reikalingas tolimesnis testavimas

X₂ – reikia palyginti prognozes su stebėtais įverčiais

Dienos šviesos lempų be namų ūkių rodiklių įtakos duomenų imtis buvo vienintelė, kai ANN PN–GRN metodo rezultatai buvo geri. Pernelyg maži duomenų rinkiniai (tik 3 stebėjimai realiam modelio potencialui ir prognozavimo gebėjimams išbandyti) galėjo lemti tai, kad neuronų tinklai nepateisino lūkesčių.

Po skaičiavimų gauti „idealūs“ rezultatai naudojant MLR bei PLS modelius buvo pašalinti iš tolimesnio tyrimo dėl prielaidos, kad naudojamos duomenų imtys buvo per trumpos, kad modeliai gerai veiktų, ir dėl to regresijos kreivės taškai sutapo su nustatytomis vertėmis. Realybėje tai mažai tikėtina.

SVM regresijos metodas gerai įvertintas tik su medicininių atliekų duomenų rinkiniais. Modelių validumui patvirtinti naudotų trijų stebėjimų užteko teigiamam SVM metodų potencialui įrodyti.

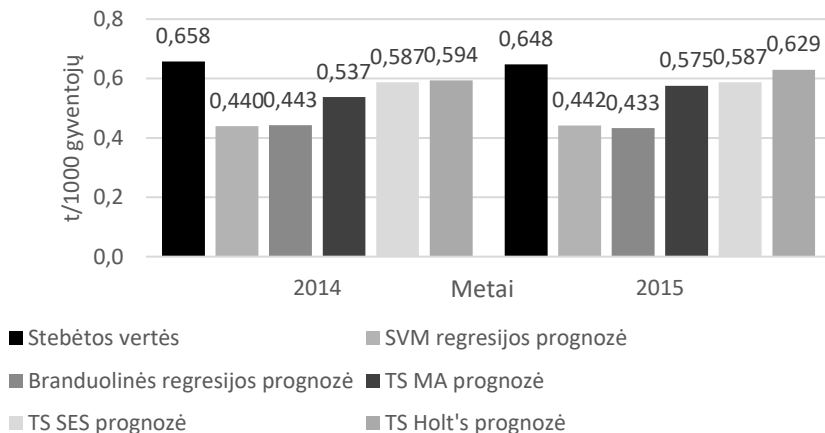
Apibendrintų adityvų metodas geriausiai veikė su dienos šviesos lempų duomenų imtimi. Lokalios regresijos rezultatai nebuvo tinkami. Glodinančių splainių metodas puikiai veikė su pavojingų medicininių ir suminių automobiline atliekų duomenų rinkiniais. Branduolinė regresija ypač gera buvo su visais duomenų rinkiniais, išskyrus sumines automobiline atliekas. Šie rezultatai yra tik teoriniai, nes modeliai nebuvo išbandyti su duomenimis, esančiais už tuo metu sudarytų duomenų imčių ribų, tačiau jei modelis teoriškai nebuvo vertintas teigiamai, praktiškai jie bus dar menkesni.

Visi laiko eilučių metodai geriausiai įvertinti su medicininių atliekų duomenų rinkiniais. Šiuo atveju rezultatai taip pat yra teoriniai.

5.2 ETAPAS: MODELIŲ TESTAVIMAS NAUDOJANT DUOMENIS UŽ MOKYMUI NAUDOTŲ DUOMENŲ IMČIŲ RIBŲ (VALIDAVIMAS)

Siekiant išbandyti sudarytų modelių validumą, 2 etapo metu buvo naudoti aktualūs 2014–2015 m. statistiniai duomenys. 2016 m. duomenys buvo įtraukti tik į galutinį tyrimo etapą ir nebuvo naudojami skaičiavimams atlikti.

5.2.1 Medicininių atliekų duomenų imtys



5.2.1.1 pav. Nustatytų ir prognozuotų susidarančių suminių medicininių atliekų kiekių palyginimas

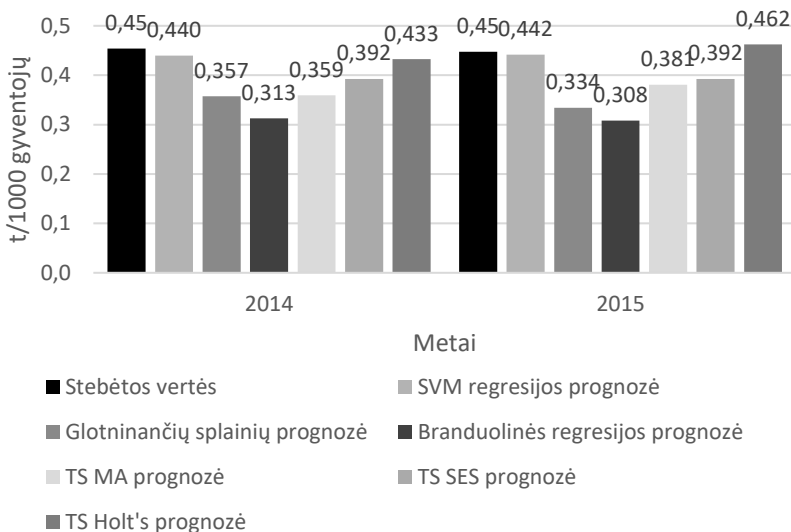
Atlikus 2 etapo bandymus su suminių medicininių atliekų duomenų rinkiniais, paaiškėjo, kad laiko eilutės turi didžiausią potencialą – visi trys metodai neviršijo kritinės 15 proc. MAPE ribos.

Kilo įtarimų dėl TS SES metodo tikslumo, kai buvo pastebėta, kad visos prognozės po 2014 m. buvo tokios pačios, kaip ir paskutinė nustatyta vertė. Paprastai taip atsitinka, kai SES metodu nenustatomos sezoninės tendencijos, o tik įvertinama, kad prognozė yra visų stebėjimų vidurkis. Taigi TS SES metodas buvo pašalintas iš tolesnių tyrimų.

Pirmajame etape SVM regresija buvo vienas iš modelių, rodančių realią prognozavimo galimybę. Tačiau šio metodo rezultatai 2 tyrimo etapo metu nebuvo vertinti teigiamai – SVM regresijos MAPE vertė labai didelė (32,48 proc.).

5.2.1.1 lentelė. Su suminių medicininių atliekų duomenų imtimi išbandytų modelių vidutinė absoliutinė procentinė paklaida (MAPE)

	SVM regresija	Branduolinė regresija	TS MA	TS SES	TS Holt
MAPE, proc.	32,48	32,93	14,77	10,00	6,29



5.2.1.2 pav. Nustatytų ir prognozuotų susidarančių pavojingų medicininių atliekų kiekių palyginimas

Trys iš septynių metodų, išbandytų su **pavojingų medicininių atliekų** duomenų imtimis, viršijo nustatytą 15 proc. MAPE ribą. SVM regresijos ir TS Holt metodai MAPE buvo labai maža – atitinkamai 2,18 proc. ir 3,96 proc. TS SES modelis buvo pašalintas iš tolesnio tyrimo dėl anksčiau nurodytų priežasčių.

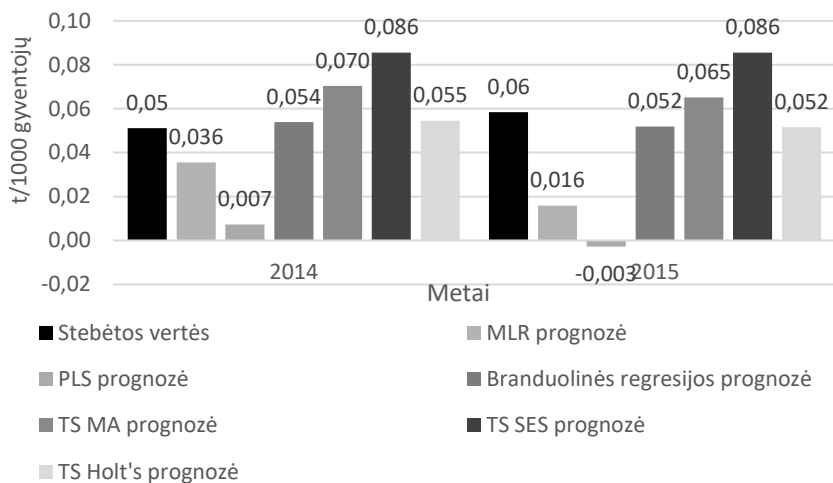
5.2.1.2 lentelė. Su pavojingų medicininių atliekų duomenų imtimi išbandytų modelių vidutinė absoliutinė procentinė paklaida (MAPE)

	SVM regresija	Glodinantys splainiai	Branduolinė regresija	TS MA	TS SES	TS Holt

MAPE, proc.	2,18	23,28	31,13	17,88	12,90	3,96
-------------	------	-------	-------	-------	-------	------

5.2.2 Dienos šviesos lempų ir atliekų, turinčių gyvsidabrio, duomenų imtis

Atlikus šio etapo tyrimus su dienos šviesos lempų atliekomis, nuspręsta nenaudoti dienos šviesos lempos atliekų susidarymo be namų ūkių suvartojimo duomenų imties tolesniuose tyrimuose ir kuriant galutinį prognozavimo modelį. Tiriant tokią patį duomenų imtį su namų ūkių išlaidų kintamaisiais naudojant pasirinktus modelius, prognozavimo rezultatai buvo daug geresni.



5.2.2.1 Nustatytų ir prognozuotų susidarančių dienos šviesos lempų atliekų palyginimas

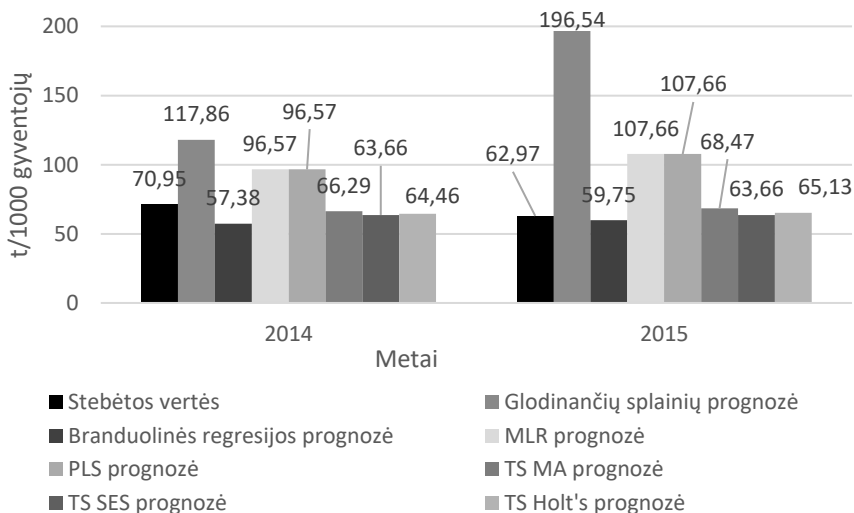
Tik du iš šešių tiriamų metodų prognozių tikslumas peržengia nustatytą 15 proc. MAPE ribą. Šie metodai yra branduolinė regresija ir Holt laiko eilutės.

5.2.2.1 lentelė. Su dienos šviesos lempų atliekų duomenų imtimi išbandytų modelių vidutinė absoliutinė procentinė paklaida (MAPE)

	MLR	PLS	Branduolinė regresija	TS MA	TS SES	TS Holt
MAPE, proc.	155,65	95,27	8,25	18,80	56,88	9,15

5.2.3 Automobilinių atliekų duomenų imtys

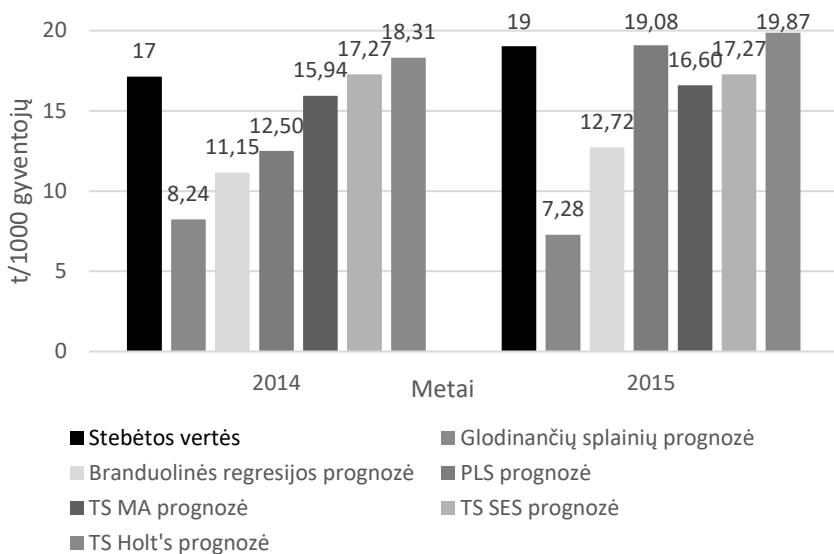
Keturių iš septynių naudojamų modelių sugeneruotos prognozės neviršija 15 proc. MAPE ribos **suminių automobilių atliekų atveju**. TS SES modelis dėl anksčiau minėtų priežasčių šalinamas iš tolimesnio tyrimo. TS MA ir TS Holt MAPE vertės panašios.



5.2.3.1 pav. Nustatytų ir prognozuotų susidarantių suminių automobilių atliekų palyginimas

5.2.3.1 lentelė. Su suminių automobilių atliekų duomenų imtimi išbandytų modelių vidutinė absoliutinė procentinė paklaida (MAPE)

	Glodinatys splainiai	Branduolinė regresija	MLR	PLS	TS MA	TS SES	TS Holt
MAPE, proc.	139,12	12,11	53,54	53,54	7,65	5,69	6,29



5.2.3.2 pav. Pastebėtų ir prognozuotų susidarantių pavojingų automobilių atliekų palyginimas

Pavojingų atliekų susidarymo tyrimo dalyje keturi iš septynių naudojamų metodų neviršijo kritinės MAPE ribos, ir nors TS SES MAPE vertė buvo mažesnė nei 15 proc., šis metodas buvo pašalintas dėl to, kad visos ateities prognozės atitinka vidutinę stebimųjų vertę. Pavojingų automobilių atliekų atvejis buvo vienintelis, kai PLS MAPE vertė buvo žemesnė nei kritinė – 13,67 proc.

5.2.3.2 lentelė. Su pavojingų automobilių atliekų duomenų imtimi išbandytų modelių vidutinė absoliutinė procentinė paklaida (MAPE)

	Glodinatys splainiai	Branduolinė regresija	MLR	PLS	TS MA	TS SES	TS Holt
MAPE, proc.	56,83	51,67	831,65	13,67	9,89	6,79	9,89

5. 3 ETAPAS: LAIKO EILUČIŲ BANDYMAI

Po atlikto 3 tyrimo etapo paaiškėjo, kad galima bandyti sudaryti hibridinius modelius, kurių paskirtis būtų suminių medicininių, suminių automobilių ir pavojingų automobilių atliekų susidarymo prognozavimas, remiantis laiko eilučių kintančio vidurkio ir Holt eksponentinio glodinimo metodais. Visais trimis

atvejais modeliams mokytį reikėtų naudoti 10 stebėjimų laiko eilutes, ir slenkančio vidurkio imtis turėtų būti lygi 3.

Holt eksponentinio glodinimo metodas gali būti naudojamas visais penkiais nagrinėtais atliekų susidarymo atvejais kaip individualus (ne hibridinis) metodas. Šiuo metodu kuriant prognozes, turėtų būti naudojama bent 8–10 stebėjimų duomenų rinkinių.

6. NAUDOTŲ IR PRITAIKYTŲ STATISTINIO MODELIAVIMO METODŲ SSGG ANALIZĖ

Siekiant atskleisti tam tikrų modelių specifkacijas, kurios turi įtakos prognozuojant konkrečių atliekų susidarymą, buvo nuspręsta atlikti trumpą stiprybių, silpnybių, galimybių ir grėsmių (SSGG) analizę.

6.1 lentelė. Pritaikytų ir sėkmingai įvertintų statistinio modeliavimo metodų stiprybių, silpnybių, galimybių ir grėsmių (SSGG) analizė

		Stiprybės	Silpnybės
Vidinės charakteristikos	PLS	PLS modelio sėkmingo pritaikymo priešzastis pavojingų automobilinių atliekų atveju yra ta, kad modeliui išbandyti naudota duomenų imtis turi daug nepriklausomų kintamųjų, kurie yra tarpusavyje kolinearūs ir susiję neapibrėžtais ryšiais. Viena iš PLS modelio funkcijų yra išskaidyti nepriklausomus kintamuosius į smulkesnes nekoreliuojančių komponentų imtis ir pritaikyti joms mažiausių dalinių kvadratų regresiją.	Modelio prognozių tikslumas sumenks, jei jas sudarant bus naudojami duomenys, kurie nepatenka į modeliui išbandyti naudotos duomenų imties ribas.

	SVM	<p>SVM modelio sėkmė, pritaikius jį pavojingoms medicininėms atliekoms, susijusi su SVM bruožu sėkmingai veikti bet kurios dimensijose ribose be išaugusių apskaičiavimo sąnaudų, jei tik atrandama tinkama branduolinė (kernelio) funkcija. Kad sėkmingai veiktu, nebūtina identifikuoti naudotos duomenų transformacijos ar kitų savybių.</p>	<p>Gali būti neatrasta tinkama branduolinė (kernelio) funkcija arba pritaikyti svorio koeficientai generuojamas prognozes per daug suvidurkins. Tokiu atveju likutinės reikšmės taps per didelės. Pagrindiniai funkcijos parametrai gali būti netinkamai parinkti.</p>
	Branduolinė regresija	<p>Branduolinės regresijos sėkmės priežastis suminių automobilių atliekų atveju yra ta, kad modelis neidentifikuoja, kokiais tarpusavio ryšiais yra susiję priklausomas ir nepriklausomi kintamieji. Šis metodas nepriklausomiems kintamiesiems pritaiko tokius svorio koeficientus, kad po atliktų skaičiavimų būtų gauta vertė, kuo artimesnė fiksuotai priklausomo kintamojo vertei.</p>	<p>Jei duomenų svyravimai nėra arba yra tik dalinai susiję tarpusavyje, tai modelis neįgalus apskaičiuoti tikslų prognozių. Pritaikyti svorio koeficientai generuojamas prognozes per daug suvidurkina.</p>
	TS Holt	<p>Ts Holt metodas veikia tiksliausiai, kai yra apdorojami duomenys su kryptinga tendencija. Modelis efektyviai išanalizuoja duomenis ir replikuoja tendencijos žingsnį naudojant dvi lyginimo ir vieną prognozavimo lygtį. Generuoja tiksliausius rezultatus, kai apdorojami duomenys su nedideliais imties svyravimais.</p>	<p>Metodas neatvaizduoja ateityje galimų duomenų svyravimų, ir ilgalaikės prognozės išsikreipia, tampa netikslios. Modelis veikia „žingsnis po žingsnio“ principu ir gali atskleisti ateities tendencijas, jei jos nenukrypsta nuo buvusių praeityje ar esamų. Duomenys turi būti išdėstyti vienodais intervalais ir chronologine tvarka.</p>
		Galimybės	Grėsmės

Modelių praktinio pritaikymo sąlygos	PLS	Metodas ypač naudingas, kai duomenų imtis turi daugiau nepriklausomų kintamųjų nei atskirų jų stebėjimų. PLS galutinis rezultatas yra tiesinė lygtis, kuri taip pat nustato nepriklausomų kintamųjų svertinius koeficientus ir reikšmingumą.	Modelis turi būti nuolat atnaujinamas. Socialiniai-ekonominiai indikatoriai generuojant prognozes turėtų būti vertinami atsargiai, įvertinant faktą, kad jie negali išeiti už modeliui mokyti naudotos duomenų imties ribų.
	SVM	Modelis geba tiksliai generuoti prognozes, kai yra naudojami tarpusavyje kolinearūs, nežinomais ryšiais susiję nepriklausomi kintamieji. R programinės įrangos aplinkoje modelis iškart pateikia modelio ir prognozių validumo rezultatus, nes pateiktą duomenų imtį padalina į mokyti ir testuoti skirtas imtis. Palengvina situacijas, kai prognozavimo procesas yra kompliktuotas dėl didelio galimai tai paveikiančių faktorių skaičiaus ir kai rankinis tinkamų faktorių parinkimas yra sunkesnis.	Nepriklausomi kintamieji turi būti parenkami siekiant išvengti papildomo duomenų „triukšmo“, modeliui nustatant panašių kintamųjų įtaką modeliavimo proceso tikslumui. Geriau parinkti daugiau kintamųjų, kurie apima skirtingus įtaką darančius aspektus.
	Branduolinė regresija	Modelis, kuris beveik visose išbandytose situacijose įvertintas teigiamai. Dėl universalių charakteristikų ir veikimo principų modelis apeina kolinearumo, duomenų buvimo už mokyti skirtos imties ribų, nežinomų ryšių tarp kintamųjų problemas. Branduolinė regresija turėtų būti pirmas išbandomas metodas modeliuojant atliekų susidarymą.	Parametriniai metodai paprastai yra efektyvesni, jei surinkti duomenys turi tam tikrą struktūrą, kuri priklauso žinomiems parametriniams modeliams.

TS Holt	Gali būti sėkmingai naudojami generuojant sezonines atliekų prognozes nacionaliniu, regioniniu ir vietiniu lygmeniu, nenaudojant ekonominių-socialinių indikatorių.	Sudarys netikslias prognozes, jei tam tikri socialiniai-ekonominiai indikatoriai darys įtaką atliekų susidarymo procesams. Dėl šios priežasties atliekų susidarymo tendencijos gali būti iškraipytos ir neapibrėžtos.
----------------	---	---

7 IŠVADOS

1. Po atliktos prieinamų mokslinių šaltinių analizės buvo nuspręsta išbandyti 6 tipų (dirbtiniai neuronų tinklai, daugialypė tiesinė regresija, dalinių mažiausių kvadratų metodas, atraminiai vektoriai, nparametrinės regresijos ir laiko eilutės) metodus, kurie galėtų taikliai prognozuoti medicininių, automobilinių ir dienos šviesos lempų bei kitų gyvsidabrio turinčių atliekų susidarymą.

2. Penki metodai (branduolinė regresija, SVM regresija, TS Holt, TS MA, PLS) pateisino lūkesčius modelių validumo bandymų etape. Suminių medicininių atliekų atveju naudojant TS MA modelį pavyko pasiekti 14,77 proc. vidutinę paklaidą, o TS Holt modelį – 6,29 proc. vidutinę paklaidą. Pavojingų medicininių atliekų duomenų imties atveju naudojant SVM modelį pavyko pasiekti 2,18 proc. paklaidą, TS Holt modelį – 3,96 proc. paklaidą. Suminių automobilinių atliekų atveju naudojant branduolinės regresijos modelį vidutinė paklaida buvo 12,11 proc., TS MA – 7,65 proc., TS Holt – 6,29 proc. Pavojingų automobilinių atliekų atveju PLS modelio vidutinė paklaida buvo 13,67 proc., TS Holt – 9,89 proc. Dienos šviesos lempų atliekų atveju geriausi rezultatai buvo naudojant branduolinės regresijos modelį (8,25 proc.) ir TS Holt (9,15 proc. vidutinė paklaida).

3. Tais atvejais, kai buvo naudojami socialiniai-ekonominiai rodikliai, geriausiai įvertinti tie statistinio modeliavimo metodai, kurių veikimas nepagrįstas ryšių tarp priklausomų ir nepriklausomų kintamųjų apibrėžimu (branduolinė regresija, SVM, PLS). TS Holt metodas veikė geriausiai tais atvejais, kai apskaičiuotos atliekų susidarymo tendencijos buvo aktualios ateityje.

4. Suminių medicininių atliekų susidarymo prognozė naudojant TS Holt modelį 2020 m. yra 0,78t/1000 gyventojų, pavojingų medicininių atliekų – 0,61

t/1000 gyventojų. TS Holt metodas pateikė prognozę, kad suminių automobilių atliekų susidarymas 2020 m. turėtų siekti 68,78 t/1000 gyventojų, pavojingų automobilių atliekų – 26,64t/1000 gyventojų. Tas pats metodas prognozuoja, kad 2020 m. turėtų būti surinkta apie 0,04t/1000 gyventojų dienos šviesos lempų ir kitų gyvsidabrio turinčių atliekų.

5. Jautrumo duomenų pokyčiai buvo fiksuojami nagrinėjant nustatytų bei prognozuotų duomenų pokyčius, nes didžiausio efektyvumo modeliai adaptavo nepriklausomų kintamųjų įtaką prognozėms taip, kad būtų pasiektas kuo didesnis efektyvumas kiekvienos atliekų susidarymo imties atžvilgiu.

8.MOKSLINIŲ PUBLIKACIJŲ DISERTACIJOS TEMA SĄRAŠAS

1. **Karpušenkaitė, A.** Ruzgas, T. Denafas, G. Forecasting medical waste generation using short and extra short datasets: Case study of Lithuania. *Waste Management & Research* 1–10, 2016. DOI: 10.1177/0734242X16628977.
2. **Karpušenkaitė, A.** Ruzgas, T. Denafas, G. Forecasting hazardous waste generation using short and extra short datasets: Case study of Lithuania. *Science – future of Lithuania*. 2016. ISSN:2029–2341, eISSN:2029–2252. vol.8, iss. 4, p.357–364.
3. **Karpušenkaitė, A.** Ruzgas, T. Denafas, G. Forecasting automotive waste generation using short and extra short datasets: Case study of Lithuania. *Environmental problems*. 2017, Vol. 2, No.1.
4. **Karpušenkaitė, Aistė;** Denafas, Gintaras; Ruzgas, Tomas. Time series based hybrid mathematical modelling method adapted to forecast automotive, medical and daylight lamps waste generation: case study of Lithuania. *Waste Management & Research*. 2018. Vol. 36(5) 454–462. DOI: 10.1177/0734242X18767308
5. **Karpušenkaitė, Aistė.** Ruzgas, Tadas. Denafas, Gintaras. *Overview of mathematical modelling methods used to forecast certain hazardous waste generation in Lithuania / Актуальные проблемы экологии: сборник научных статей по материалам XII Международной научно–практической конференции, (Гродно, 4–6 октября 2017 г.) / Гродненский государственный университет. Гродно : ЮрСаПринт, 2017. p. 132–134. SBN: 9789857134426*
6. **Karpušenkaitė, Aistė.** Ruzgas, Tomas. Denafas, Gintaras. *Forecasting daylight lamps waste and waste that has mercury in its composition generation using extra short data sets: case study of Lithuania*. Linnaeus Eco–Tech 2016: proceedings of the international conference on natural sciences and technologies for Waste and wastewater treatment remediation emissions related to climate environmental and economic

effects: 10th international conference on the establishment of cooperation between companies and institutions in the Nordic countries, the Baltic Sea region and the world, November 21–23, 2016, Kalmar, Sweden. [Växjö]: [Linnaeus University], 2016. p. 1–8. ISBN: 9789188357410.

9. TRUMPOS ŽINIOS APIE AUTORIŲ

Aistė Karpušenkaitė gimė 1988 m. lapkričio 26 d. 2008 m. baigė Šiaulių Lieporių vidurinę mokyklą.

2008–2012 m. mokėsi ir baigė Aplinkos ir profesinės saugos bakalauro studijas Šiaulių universiteto Technologijos fakultete.

2012–2014 m. mokėsi ir baigė Aplinkos apsaugos vadybos ir švaresnės gamybos magistro studijas Kauno technologijos universiteto Aplinkos inžinerijos institute.

2014–2018 m. mokėsi doktorantūros studijų programoje „Aplinkos inžinerija ir kraštovarkla“ Kauno technologijos universiteto Cheminės technologijos fakultete.

SUMMARY

Proper hazardous waste management and treatment is a worrying matter in the whole waste management system. Especially in countries such as Lithuania, where even the normal solid waste disposal into landfills is quite different from the established ones by the European Union and is considered to be flawed due to ineffective sorting in primal waste generation sources and inefficient waste collection and identification system. Due to these reasons major part of hazardous waste from households reach their end of life in landfills as a constituent of solid waste. Furthermore, unofficially there is an illegal practice in industry, which falsely identifies hazardous waste as other types, usually less toxic or cheaper to manage, waste. It is thought that if hazardous waste collection, management, identification, and control system would work in its full efficiency, there would be 1/3 more waste than it is today. Scientists that work in this field often state there is a lack of information about the factors that could improve the system and about the current situation characteristics and the reason why desired results are not always achieved.

The possibility of making accurate and reasonable forecasts on various types and total hazardous waste generation is an important aspect in making waste management decisions in short and long terms. Predictions based on openly accessible socio-economic factors provide necessary data for various interested parties from public and business sectors. Such knowledge can influence the

choosing and development of effective hazardous waste utilization methods or the establishment of certain facilities.

The overview of previously conducted research on waste matter in general showed that various mathematical forecasting methods were used to predict solid waste generation (SWG). Only a few published research articles that were able to be found using a wide access of scientific journals and subscribed scientific databases uncovered successful application of mathematical prognostic methods applicable for medical waste and barely any for other types of hazardous waste.

Medical waste management system in Lithuania is flawed as medical waste collection and separation systems are not efficient in hospital and clinics, medical waste collection system from households is barely established and publicly announced or talked about. Majority of Lithuanians don't know that they should and could bring used syringes, medicine bottles or other infection and hazardous medical waste to pharmacies or hospitals. There are known cases when pharmacists decline to accept such waste even though they should and probably it is because there is no well-organized collection system from establishments other than hospitals and clinics. Because of these reasons medical, including hazardous, waste often ends up in municipal solid waste flow and later reaches landfills where they are not always sorted out.

Automotive waste generation is also growing every year and along with it grows the danger that is caused by hazardous components in vehicles which are still not always treated responsibly and end up in the city landfill. The need to know how and which socio-economic factors influence automotive waste generation is growing bigger when not only Lithuania's but also other European Union countries governments and waste managers want to effectively control and manage total and hazardous waste collection, management, and treatment system. Forecasting would make it easier to prepare necessary capacities to ensure that such public and private sectors plans would reach its maximum goal.

Considering fluorescent tubes and other mercury-containing generation data it is clear that amounts of these waste declared by Environmental protection agency mainly comes from industry, business and service sectors. Lamp waste that is generated by households often ends up in municipal solid waste flow if not being separated in larger quantities in special collection sites. Since mechanical biological treatment facilities are only beginning to be established in Lithuania it is hard to say what the real amount of fluorescent tubes and other mercury-containing waste could be.

Research object – medical waste, automotive waste, fluorescent tubes, and other mercury-containing waste generation.

The aim of the research – the development of statistical model(s) which would provide an accurate (error rate <15 proc.) forecast for chosen waste generation.

Research tasks:

1. Analyse and compare waste generation prediction methods and practices found in accessible literature, determine the most efficient ones;
2. Analyse models training and testing (validation) stages results and pick out statistical models that show most promising results in certain type waste forecasting;
3. Forecast potential medical waste, automotive waste, fluorescent tubes, and other mercury-containing waste generation up till 2020.
4. Conduct a sensitivity analysis on collected and forecasted waste generation data, compare received results between observed and forecasted data.

Methodology and approach

Annual data about certain types of waste generation in Lithuania during 2004–2015 was gathered and analysed. Datasets including waste generation and it influencing socio-economic factors were created. These data sets were used to test 13 different methods from 6 types (artificial neural networks, multilinear regression, partial least squares, support vector machine, non-parametrical regression, time series) of statistical modelling methodologies. Five models passed the verification stage and showed the best results.

Scientific novelty and practical value

This research compared many different statistical modelling methods and determined which models with their specific features would work best with medical, automotive and fluorescent tubes and other mercury-containing waste generation data. Approaches to forecast with and without influencing factors were tested and revealed that in certain situations different models have advantages. Research of this broad scale and including different types of hazardous waste yet searching for similarities between tendencies in their generation was done one of the first times and has laid foundations for further in-depth research exploring possibilities to develop more unique forecasting models for hazardous or other specific waste generation.

Hypothesis

Statistical modelling methods can be successfully and effectively applied for forecasting when training them on short datasets made from hazardous and other certain types of waste generation data.

Structure and outline of the dissertation

This dissertation is divided into these parts: introduction, short introduction to waste management sector in Lithuania, literature analysis, methodology, certain types of waste generation and dependability analysis, results of stage 1 – generated forecasts and models performance evaluation, results of stage 2 – testing models outside training data set limits, results of stage 3 – testing flexibility of time series methods, conclusions, list of 73 references, and 8 annexes.

Conclusions

1. Conducted literature analysis helped to determine that 6 types (artificial neural networks, multilinear regression, partial least squares, support vector machine, non-parametrical regression, time series) of methods could show high potential while forecasting medical waste, automotive waste and fluorescent tubes waste generation in Lithuania's case.

2. Five models (kernel regression, SVM regression, TS Holt, TS MA, PLS) passed model validation stage and showed best results. In total medical waste generation case TS Holt showed 6.29 proc. MAPE. Hazardous medical waste generation case – SVM demonstrated 2.18 proc. MAPE, TS Holt – 3.96 proc. MAPE. In total automotive waste generation case – kernel regression showed 12.11 proc. MAPE, Holts – 6.29 proc. MAPE. In hazardous automotive waste generation case – PLS showed 13.67 proc. MAPE, TS Holt – 9.89 proc. MAPE. In fluorescent tubes and other mercury-containing waste generation case – kernel regression showed 8.25 proc., TS Holt – 9.15 proc. MAPE.

3. Models which were able to process data and generate forecasts without presuming or determining relationships between dependent and independent variables performed best (Kernel regression, PLS, SVM) when socio-economic indicators were used. TS Holts produces accurate forecasts when calculated trends were valid for the future waste generation.

4. It was estimated, that in 2020 total medical waste generation will be 0.78 t/1000 residents, hazardous medical waste – 0,61 t/1000. According to TS Holt forecasts there should be 68.78 t/1000 residents of total automotive waste and 26,64 t/1000 residents of hazardous automotive waste. TS Holt forecasted that fluorescent tubes and other mercury-contained waste generation should be 0.04 t/1000 residents in 2020.

5. Data sensitivity analysis revealed that changes between independent and dependant variables sensitivity were noticed as a model that demonstrated highest efficiency enhanced variable's sensitivity to develop better and more efficient models structure for accurate forecasting.

UDK – 628.4.045/.046:519.2 (474.5)(043.3)

SL344. 2019–05–31, 2,5 leidyb. apsk. l. Tiražas 50 egz.

Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas