

STATISTINIŲ METODŲ TAIKYMAS VERTINANT KLAIPĖDOS MIESTO BUTŲ KAINAS

Roberta Navickaitė

VĮ Registrų centras. Adresas: Vinco Kudirkos g. 18-3, LT-03105 Vilnius, Lietuva
E. paštas: roberta.navickaite@registrucentras.lt

Gauta: 2014 m. liepa Pataisyta: 2014 m. spalio Paskelbta: 2014 m. lapkritis

Santrauka. Straipsnyje nagrinėjamas netiesinės regresijos modelis, apibendrintasis tiesinis modelis ir apibendrintasis adityvus modelis (pusiau parametrinis regresijos modelis), sudarant nekilnojamojo turto vertinimo modelį. Šie modeliai pritaikyti Klaipėdos miesto butų pirkimo ir pardavimo sandorių duomenims. Tyrimo tikslas – sudaryti nekilnojamojo turto verčių regresijos modelius, taikant įvairius statistinius metodus, ir palyginti gautus modelius tarpusavyje. Šiame straipsnyje išnagrinėti praktiniai regresijos modelių sudarymo aspektai bei pateikiamos tyrimo išvados.

Reikšminiai žodžiai: apibendrintasis adityvus modelis, apibendrintasis tiesinis modelis, butai, netiesinės regresijos modelis, pusiau parametrinis regresijos modelis.

1. Įvadas

Nekilnojamoju turtu vadinamas žemės sklypas ir su juo susiję daiktai, kurie negali būti perkelti iš vienos vietos į kitą nepakeitus jų paskirties ir iš esmės nesumažinus jų vertės [13]. Nekilnojamojo turto vertinimo taisyklės vertinimo modelį apibrėžia taip: „Nekilnojamojo turto vertinimo modelis – matematinė formulė, taikoma nekilnojamojo turto vidutinei rinkos vertei <...> apskaičiuoti, atsižvelgiant į nekilnojamojo turto kadastro rodiklius“ [18]. Manytina, kad tinkamiausias būdas įvertinti visos šalies nekilnojamojo turto vertinimo modelius naudojantis nekilnojamojo turto sandorių duomenų baze. Nekilnojamojo turto vertinimas yra svarbus ir aktualus ne tik įsigyjant nekilnojamojo turto, bet ir apmokestinamo turto mokesčių verčių apskaičiavimui, valstybės valdomo turto pardavimo ar nuomos kainų nustatymui, mokesčių turto paveldėjimo, dovanojimo atvejais vertės apskaičiavimui, notarinių paslaugų kainų, tvirtinant turto perleidimo sandorius, nustatymui, skiriant socialinę paramą ir kitoms reikmėms [10, 24].

Nekilnojamajam turtui vertinti pasaulyje plačiai naudojamas nekilnojamojo turto vertinimo modelis, sudarytas taikant tiesinės regresijos modelį. Tačiau Tarptautinės turto vertintojų mokesčiams asociacijos (ang. *International Association of Assessing Officers* – IAAO) buvo pasiūlytas nekilnojamojo turto vertinimo modelis, sudarytas taikant netiesinės regresijos modelį [7]. Remiantis siūlomu nekilnojamojo turto vertinimo modeliu [7], šiame darbe sudaryti nekilnojamojo turto vertinimo modeliai Lietuvos sąlygomis, taikant netiesinės regresijos modelį [20], apibendrintąjį tiesinį modelį [2, 6, 19, 23] ir apibendrintąjį adityvųjį modelį [21, 26]. Modeliai taikomi Klaipėdos miesto butų pirkimo ir pardavimo sandorių 3 metų laikotarpio duomenims (2009 m. sausio mėn.–2012 m. kovo mėn.).

Darbo skaičiavimams atlikti bei modeliams sudaryti naudojami statistiniai paketai R [22] ir NCSS [15].

Šio darbo tikslas – sudaryti nekilnojamojo turto verčių regresijos modelius, taikant įvairius statistinius metodus, ir palyginti gautus modelius tarpusavyje. Su straipsnio tikslu susieti šie pagalbiniai uždaviniai: atrinkti reikšmingus kintamuosius, kurie naudojami nekilnojamojo turto vertinimo modeliuose, įtraukti į nekilnojamojo turto vertinimo modelius geografines koordinates, kurios nusako nekilnojamojo turto objekto vietos įtaką, sudaryti nekilnojamojo turto vertinimo modelius, taikant netiesinės regresijos modelius, apibendrintuosius tiesinius modelius ir apibendrintuosius adityviuosius modelius, palyginti gautus modelius tarpusavyje ir rasti turimiems duomenims tinkamiausią modelį. Publikacijos naujumas sietinas su nekilnojamojo turto vertinimo modelių tobulinimu Lietuvoje bei tikslesnės nekilnojamojo turto vietos nusakymas įtraukiant geografines koordinates kaip kintamuosius į vertinimo modelius.

2. Butų pirkimo ir pardavimo sandorių duomenų analizė

Statistiniais metodais vertinant nekilnojamąjį turtą, naudojami Klaipėdos miesto butų pirkimo ir pardavimo sandorių duomenys (4427 stebiniai) nuo 2009 m. sausio mėn. iki 2012 m. kovo mėn. imtinai. Šie duomenys gauti iš VĮ Registrų centro nekilnojamojo turto sandorių duomenų bazės.

Šiame tyrime nagrinėjami tokie pagrindiniai buto požymiai:

- sandorio suma,
- 1 kv. m kaina,
- buto plotas,
- zona (administracinės teritorijos suskirstymas į rajonus),
- namo statybos pabaigos metai,
- sienų medžiaga,
- buto kambarių skaičius,
- aukštas, kuriame yra butas,
- dujų įranga name,
- rūšio priklausinys,
- šildymas,
- kanalizacija,
- X ir Y koordinatės (Klaipėdos miesto butų geografinės koordinatės – Žemės paviršiaus taško (namo, kuriame yra butas) geografinė platumo ir geografinė ilguma).

Toliau trumpai apžvelgiami buto požymiai ir taikyti įvairūs statistiniai metodai [5], vertinant Klaipėdos miesto butų kainas.

Priklausomas kintamasis, kurio pagrindu įvertinama nekilnojamojo turto kaina, yra sandorio suma. Remiantis turimais duomenimis gauta, kad Klaipėdos mieste butų kainos svyruoja nuo 15 000 Lt iki 750 000 Lt, o vidutinė sandorio suma yra 138 000 Lt (žr. 1 lentelę). Kitas nekilnojamojo turto kainos nustatymo rodiklis – 1 kv. m kaina. Šis kintamasis gaunamas sandorio sumą padalijus iš objekto ploto. Galima sakyti, jog 1 kv. m kaina ir sandorio suma yra lygiavertės savybės, tačiau vertinant ar lyginant sandorius tarpusavyje, palyginimui parenkamas 1 kv. m kainos kintamasis. Iš statistinių rodiklių 1 lentelės matyti, kad 1 kv. m kainos kinta nuo 1 014 Lt iki 10 060 Lt, o vidutinė 1 kv. m kaina yra 2 690 Lt.

Priklausomai nuo objekto ploto dydžio kinta ir buto kaina. Nekilnojamojo turto bendrasis plotas apibrėžiamas kaip buto kapitalinių sienų ribų, pastogių, rūsių (pusrūsių), uždarytų lodžijų, esančių patalpų plotų suma [16]. Šis rodiklis turi didžiausią įtaką vertinimo modeliui. Iš 1 lentelės pastebėtina, jog vidutinis nagrinėjamų butų plotas yra apie 51 kv. m, tokio dydžio butus galima vadinti standartiniais.

1 lentelė. Kintamųjų (sandorio sumos, 1 kv. m kainos, objekto ploto) statistiniai rodikliai

Statistinis rodiklis	Sandorio suma	1 kv. m kaina	Objekto plotas
Minimali reikšmė	15 000	1 014	10,55
Pirmasis kvartilis	92 000	2 165	36,58
Mediana	125 000	2 561	48,09
Vidurkis	138 000	2 690	51,07
Trečiasis kvartilis	160 000	3 026	63,02
Maksimali reikšmė	750 000	10 060	255,40
Standartinis nuokrypis	73 309,94	828,31	18,88

Vienas iš svarbiausių veiksnių, vertinant nekilnojamąjį turtą, yra vietos įtaka. Pasinaudojus VĮ Registrų centro Klaipėdos miesto savivaldybės teritorijos nekilnojamojo turto vertinimo dokumentais [25], nustatomos verčių zonos. Verčių zona – tai teritorija, kurioje žemės sklypų charakteristikos ir kainos yra panašios [1] (toliau – zona). Remiantis Klaipėdos miesto savivaldybės teritorijos nekilnojamojo turto verčių zonų žemėlapiu [25], nustatoma, kurioje miesto dalyje yra vertinamas objektas. Klaipėdos miesto savivaldybėje iš viso yra 32 zonos, tačiau nagrinėjami butų paskirties

objektai Klaipėdos mieste yra išsidėstę 16-oje zonų. Įtraukiant 16 zonų kaip atskirus kintamuosius į regresijos modelį, pats modelis gali tapti sudėtingas, todėl turimos zonos, kurios yra panašios 1 kv. m kainos atžvilgiu, suskirstytos į tris grupes ir kiekviena zonų grupė užkoduota pseudokintamaisiais, t. y. priskiriami dvireikšmiai kintamieji.

Namo statybos pabaigos metai – vienas iš reikšmingų rodiklių. Nekilnojamojo turto objekto kaina priklauso nuo metų, kuriais buvo jis pastatytas. Tačiau senos statybos nekilnojamojo turto kainai gali turėti įtakos vieta (zona). Yra žinoma, kad miesto centre ar senamiestyje nekilnojamojo turto objektai yra brangiausi, nors šioje teritorijoje nekilnojamojo turto objektai yra senos statybos. Namų statybos pabaigos metų kintamasis pertvarkytas į tris senumo grupes (namų statybos pabaigos metai iki 1994 m., nuo 1995 iki 1974 m. ir nuo 1975 m.) bei atitinkamai priskirti dvireikšmiai kintamieji.

Kitas nagrinėjamas kintamasis – sienų medžiaga. Turint 4 pagrindinius sienų medžiagų tipus (blokeliai, gelžbetonio plokštės, monolitinis gelžbetonis ir plytos), įvertinta, jog vidutinė 1 kv. m kaina atitinkamai yra 3674 Lt, 2411 Lt, 3209 Lt ir 2837 Lt. Pastebėtina, kad didžiąją nagrinėjamų duomenų dalį sudaro daugiabučiai, kurių sienų medžiagos yra plytos ir gelžbetonio plokštės. Tačiau daugiabučių, kurie pastatyti iš blokelių, vidutinė 1 kv. m kaina, palyginti su kitais sienų medžiagų tipais, yra didžiausia. Mažiausia vidutinė 1 kv. m kaina yra pastatuose, kurių sienų medžiaga yra gelžbetonio plokštės. Sienų medžiagų tipai užkoduoti pseudokintamaisiais.

Nors kambarių skaičiaus bute kintamasis koreliuoja su objekto ploto dydžiu, tačiau yra tokių atvejų, kai vienodo ploto dydžio butai gali turėti ne tokį pat kambarių skaičių butuose. Tiek su objekto plotu, tiek su kambarių skaičiumi pastebėtina tendencija, jog 4 ir daugiau kambarių butai nėra paklausūs, t. y. sandorių skaičius yra nepakankamas, kai kambarių skaičius yra didesnis nei 4, palyginti su kitomis grupėmis. Todėl didesnio kambarių skaičiaus butai sujungiami į vieną grupę. Kambarių skaičiaus kintamasis suskirstomas į 4 grupes (1 kambario, 2 kambarių, 3 kambarių bei 4 ir daugiau kambarių butai) ir atitinkamai priskiriami dvireikšmiai kintamieji.

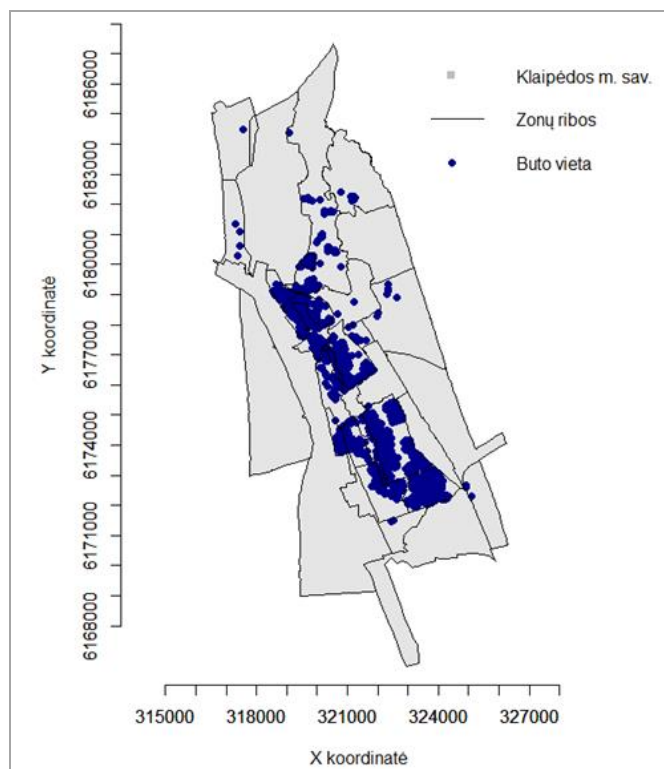
Kitas veiksnys, darantis įtaką objekto kainai – aukštas, kuriame yra butas. Iš vertinimo praktikos išsiskiria keli aukštai, kurie turi didesnę įtaką buto kainai. Tai pirmas ir viršutinis aukštai bei šiems kintamiesiems priskiriami dvireikšmiai kintamieji.

Išnagrinėjus namų dujų įrangą, gauta, kad 1 kv. m kaina yra mažesnė, kai dujos yra įrengtos name nei neturint dujų įrangos. Buto kainai įtakos turi ir rūšio priklausinys. Šiuo atveju 1 kv. m kainų vidurkių skirtumas yra 152 Lt, nesant ir esant rūšio patalpai. Dujų įrangos ir rūšio kintamieji užkoduojami pseudokintamaisiais.

Šiuo laikotarpiu gana svarbus veiksnys – šildymas. Išaugusios šildymo kainos, priklausomai nuo šildymo tipo (centrinis, krosninis ir vietinis centrinis), buto kainai taip pat turi įtakos. Gauta, jog 1 kv. m kaina yra mažesnė centrinio šildymo atveju nei krosninio ar vietinio centrinio. Yra žinoma, kad didžiąją sandorių dalį sudaro butai, kuriuose yra centrinis šildymas. Šio kintamojo reikšmėms priskiriami pseudokintamieji.

Dažniausiai butuose yra kanalizacija (nuotekų šalinimas), tačiau yra ir tokių objektų, kuriuose kanalizacijos nėra. Nagrinėjamu atveju yra 12 butų, kuriuose nėra kanalizacijos ir 15 butų, kuriuose yra vietinis nuotekų šalinimas. Taip pat atkreiptinas dėmesys, jog beveik visus turimus duomenis sudaro komunalinis nuotekų šalinimas, todėl dėl mažo sandorių kiekio šis kintamasis nebuvo nenagrinėjamas ir į modelius neįtrauktas.

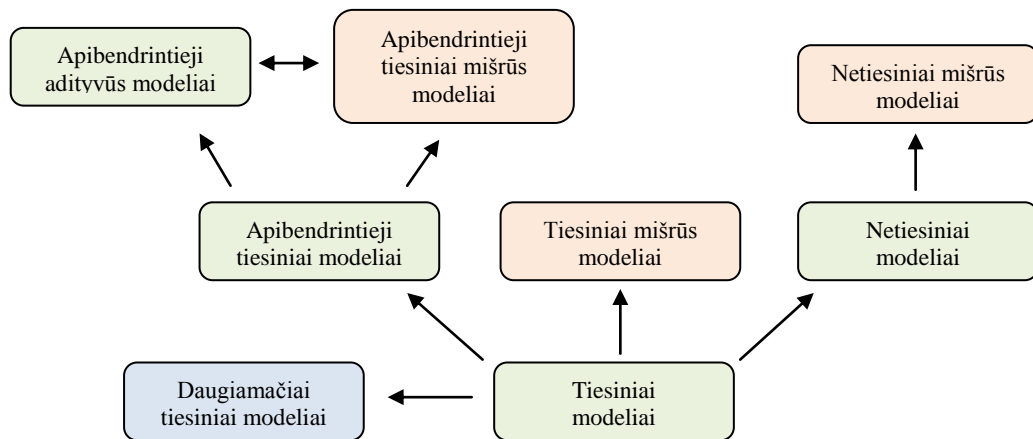
Nekilnojamojo turto objekto vietos nustatymas apibrėžiamas adresu. Su objekto adresu taip pat susietos ir geografinės koordinatės (X ir Y koordinatės). Tai koordinatės stačiakampėje koordinačių sistemoje, kurioje bet kurio taško padėtis Žemės paviršiuje nusakoma geografinės platumos ir geografinės ilgumos kampais. Jos naudojamos taškų padėčiai Žemės paviršiuje nustatyti ir nurodyti [12]. Turint taškus – geografinius duomenis, kurie išreiškiami kaip taškai plokštumoje, galima atlikti geografinių duomenų analizę [8, 9, 13]. Nekilnojamojo turto objekto pirkimo ir pardavimo sandorius išreiškus kaip taškus plokštumoje, Klaipėdos miesto butų pirkimo ir pardavimo sandoriai pavaizduojami pateikiant geografines (X ir Y) koordinatas (žr. 1 pav.).



1 pav. Sandoriuose dalyvaujančių butų išsidėstymas Klaipėdos miesto savivaldybėje

3. Statistinių modelių taikymas butų kainoms vertinti

Nekilnojamajam turtui vertinti gali būti taikomi įvairūs nekilnojamojo turto vertinimo modeliai. Pats paprasčiausias ir dažniausiai naudojamas modelis – tiesinės regresijos modelis [3, 4], tačiau nekilnojamojo turto vertinimo praktikoje tiesinės regresijos modelis ne visada pasiteisindavo. Tai buvo nagrinėta baigiamajame bakalauro darbe „Butų vertės prognozavimas, naudojant tiesinės ir netiesinės regresijos modelius“ [17] ir remiantis Tarptautinės turto vertintojų mokesčiams asociacijos pateiktais vertinimo modeliais [7]. Šiame darbe ieškota kitų metodų, tinkančių nekilnojamojo turto objektų kainoms prognozuoti. Remiantis regresijos modeliavimo metodų schema (žr. 2 pav.) [11], kurioje vaizduojamas skirtingų statistinių modelių ir jų kilmės planas, šiame tyrime turimiems Klaipėdos miesto butų pirkimo ir pardavimo sandoriams pasirinkti ir taikyti tokie regresijos modeliai – netiesinės regresijos modelis, apibendrintasis tiesinis modelis ir apibendrintasis adityvus modelis. Apibendrintasis adityvus modelis suprantamas kaip pusiau parametrinis modelis, nes jo bendra struktūra susideda iš parametrinių ir nparametrinių komponentų. Pateiktoje schemoje nagrinėti regresijos modeliai išskirti žalia spalva. Tęsiant šį tyrimą, būtų galima panagrinėti ir kitus regresijos modelius, kurie pateikti 2 paveikslėlyje.

2 pav. Regresijos modeliavimo metodų planas¹

Tarptautinė turto vertintojų mokesčiams asociacija (ang. *International Association of Assessing Officers – IAAO*) pateikė bendrą netiesinį nekilnojamojo turto vertinimo modelį [0]:

$$Y = \prod_{j=1}^k \beta_j^{x_j} \left(\left(\prod_{j=k+1}^l \beta_j^{x_j} \cdot \sum_{j=l+1}^m \beta_j x_j \right) + \left(\prod_{j=m+1}^r \beta_j^{x_j} \cdot \sum_{j=r+1}^p \beta_j x_j \right) + \sum_{j=p+1}^q \beta_j x_j \right) + \varepsilon, \quad (1)$$

čia: Y – priklausomas kintamasis (nekilnojamojo turto kaina), $Y = (Y_1, Y_2, \dots, Y_n)'$; n – imties dydis; $(x_1, x_2, \dots, x_q)'$ – nepriklausomi kintamieji, kai $j = 1, 2, \dots, k$, tai x_j – kategoriniai kintamieji, kurie atspindi viso vertinimo objekto faktorius, būdingi bet kuriam vertinamam objektui, kai $j = k+1, k+2, \dots, l$, tai x_j – kategoriniai kintamieji ir kai $j = l+1, l+2, \dots, m$, tai x_j – kiekybiniai kintamieji, kurie apibrėžia tik statinių vertinimo charakteristikas, kai $j = m+1, m+2, \dots, r$, tai x_j – kategoriniai kintamieji ir kai $j = r+1, r+2, \dots, p$, tai x_j – kiekybiniai kintamieji, kurie apibrėžia tik žemei būdingas charakteristikas, kai $j = p+1, p+2, \dots, q$, tai x_j – kiekybiniai kintamieji, kurie apibrėžia kitas vertinimui būdingas charakteristikas; $\beta_1, \beta_2, \dots, \beta_q$ – nežinomi modelio parametrai; ε – modelio paklaida, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$, čia ε_i – sudaro baltąjį triukšmą.

Į modelį gali įeiti santykiniai, dvireikšmiai ir kiekybiniai nepriklausomi kintamieji. Kiekybiniai kintamieji dažniausiai išreiškiami kaip ploto dydžiai, o santykiniai ar dvireikšmiai kintamieji daro įtaką ploto kintamajam taip sumažindami kainą arba ją padidindami. Pavyzdžiui, norint įvertinti visą namų valdą, į modelį įtraukiami visų pastatų ir žemės vertinimo požymiai. Tačiau Lietuvoje naudojamas supaprastintas (1) modelio pavidalas: vertinant statinius, naudojama tik ta modelio dalis, kurioje yra tik statinius apibūdinantys kintamieji, o vertinant žemę – tik žemę apibūdinantys kintamieji. Taigi, Lietuvoje statiniai ir žemė vertinami atskirai.

Šiame darbe nagrinėjami butų pirkimo ir pardavimo sandoriai. Butai suprantami kaip patalpos pastate ar pastato dalis ir priskiriami prie statinių. Tuomet bendra matematinė forma nekilnojamojamam turtui vertinti iš (1) modelio įgyja tokį pavidalą:

$$Y_i = \prod_{j=1}^{m_1} \beta_j^{x_{ij}} \cdot \sum_{j=m_1+1}^{m_1+m_2} \beta_j x_{ij} + \varepsilon_i, \quad (2)$$

¹ Kuhnert, P.; Venables, B. 2005: *An Introduction to R: Software for Statistical Modelling & Computing*. Australia: CSIRO Mathematical and Information Sciences. 141 psl.

čia: Y_i – priklausomas kintamasis, t. y. i -ojo nekilnojamojo turto pirkimo ir pardavimo kaina (sandorio suma); $x_{i1}, x_{i2}, \dots, x_{i, m_1+m_2}$ – nepriklausomi kintamieji, kai $j = 1, 2, \dots, m_1$, tai x_{ij} išreiškiami kaip dvireišmiai kintamieji, kai $j = m_1+1, m_1+2, \dots, m_1 + m_2$, tai x_{ij} išreiškiami kaip kiekybiniai kintamieji; $\beta_1, \beta_2, \dots, \beta_{m_1+m_2}$ – nežinomi modelio parametrai; ε_i – sudaro baltąjį triukšmą.

Ši (2) forma suprantama kaip netiesinės regresijos modelis su žinoma modelio struktūra, t. y. ryšiu tarp priklausomo ir nepriklausomų kintamųjų. (2) netiesinės regresijos modelis taikytas praktiškai turimoms Klaipėdos miesto butų kainoms prognozuoti. Taigi užkoduoti Klaipėdos miesto butų kintamieji įrašomi į (2) modelį ir NCSS statistine programa gauti modelio parametrų įverčiai pateikiami 3 lentelėje. Gautas modelis toliau žymimas ir naudojamas kaip 1 modelis.

Pastebėtina, kad šį modelį galima supaprastinti, kai naudojamas tik vienas kiekybinis kintamasis (objekto plotas), t. y. (2) modelio išraiškai pritaikoma logaritmo transformacija:

$$\ln(E(Y|x)) = \ln\left(\prod_{j=1}^{m_1} \beta_j^{x_j} \cdot \beta_{m_1+1} x_{m_1+1}\right).$$

Pritaikant logaritmo savybę, gaunama:

$$\ln(E(Y|x)) = \sum_{j=1}^{m_1} x_j \cdot \ln \beta_j + \ln \beta_{m_1+1} + \ln x_{m_1+1}.$$

Iš (2) netiesinės regresijos modelio struktūros gautas tiesinės regresijos modelis. Pažymėdami logaritmuotus kintamuosius naujais kintamaisiais:

$$Z = \ln(E(Y|x)), \gamma_j = \ln \beta_j, j = 1, 2, \dots, m_1, \text{ ir } \gamma_0 = \ln \beta_{m_1+1},$$

gaunamas toks modelis:

$$Z = \gamma_0 + \sum_{j=1}^{m_1} \gamma_j x_j + \ln x_{m_1+1} + \varepsilon,$$

čia x_j , kai $j = 1, 2, \dots, m_1$, žymi kategorinius nepriklausomus kintamuosius, o x_{m_1+1} – kiekybinį kintamąjį (objekto plotą); ε – modelio paklaida, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$, kur ε_i sudaro baltąjį triukšmą.

Šiuo atveju gaunama tiesinė kintamųjų priklausomybė. Norint gautą modelį palyginti su pirmuoju modeliu (netiesinės regresijos modeliu – 1 modeliu), reikia modelį atversti į pradinę išraišką – nelogaritmuotą struktūrą, panaudojant eksponentės transformaciją. R programos pagalba gauti modelio parametrų įverčiai pateikiami 3 lentelėje. Šis modelis toliau vadinamas 2 modeliu.

Vienas iš populiariausių metodų, parenkant priklausomo kintamojo transformaciją regresijoje, yra Bokso ir Kokso (Box ir Cox (1964)) transformacijos taikymas, kuris kartais modelio liekanų skirstinį padaro normalųjį. Šis metodas siūlo vietoje y nagrinėti naują kintamąjį $g_\lambda(y)$:

$$g_\lambda(y) = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \text{jei } \lambda \neq 0, \\ \log(y), & \text{jei } \lambda = 0. \end{cases} \quad (3)$$

Pertvarkomas priklausomas kintamasis y , taikant įvertintą parametro λ reikšmę ($\lambda = 0,1735$), ir sudaromas naujas tiesinės regresijos modelis. Modelis toliau žymimas 3 modeliu ir jo gauti parametrų įverčiai pateikiami 3 lentelėje.

Kitas nagrinėjamas atvejis – logaritmo taikymas priklausomam kintamajam. Šiuo atveju pasirinktas apibendrintasis tiesinis modelis ir pasirinkta jungties funkcija – logaritmė. Bendrasis apibendrintasis tiesinis modelis – tai tiesinis modelis, apibūdinantis priklausomo kintamojo Y vidurkio $\mu = EY$ funkciją:

$$g(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im}, i = 1, 2, \dots, n. \quad (4)$$

Funkcija $g(\cdot)$ vadinama jungties funkcija arba jungtimi. Ji parenkama atsižvelgiant į tiriamą problemą. Tiesinė modelio dalis su nežinomais parametrais $\beta_j, j = 1, 2, \dots, m$, ir nepriklausomais kintamaisiais $X_j, j = 1, 2, \dots, m$, vadinama modelio prediktoriumi bei žymima

$$\eta_i(\beta) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im}, i = 1, 2, \dots, n. \quad (5)$$

Modelis turi tenkinti tokius reikalavimus:

1. $\mu_i = EY_i$, čia Y_i skirstinys turi priklausyti eksponentinių skirstinių šeimai.
2. Jungties funkcija $g(\cdot)$ turi būti tolydi ir turėti atvirkštinę funkciją g^{-1} , t. y. $\mu_i = g^{-1}(\eta_i(\beta))$.

Gauti modelio parametrų įverčiai pateikti 3 lentelėje. Gautas modelis toliau suprantamas ir naudojamas kaip 4 modelis.

Turimiems duomenims taikomas apibendrintasis adityvus modelis (pusiau parametrinis regresijos modelis), kurį sudaro apibendrintasis tiesinis modelis pridėdamas neparimetrinius vieno kintamojo modelius. Apibendrintasis adityvus modelis – tai apibendrintasis tiesinis modelis, sudarytas iš dviejų dalių sumos: tiesinio prediktoriaus ir neparimetrinių kai kurių nepriklausomų kintamųjų glodžių funkcijų.

Bendra modelio struktūra:

$$g(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im} + s_1(z_{i1}) + s_2(z_{i2}) + \dots + s_p(z_{ip}), i = 1, 2, \dots, n, \quad (6)$$

čia $\mu_i = EY_i$ ir Y_i skirstinys priklauso eksponentinių skirstinių šeimai. Čia Y_i – priklausomas kintamasis; X_{ij} , $j = 1, 2, \dots, m$, – nepriklausomas kintamasis; β_{ij} , $j = 1, 2, \dots, m$, – nežinomi modelio parametrai; s_k , $k = 1, 2, \dots, p$, – regresorių glodžios funkcijos; z_{ik} , $k = 1, 2, \dots, p$, – nepriklausomi tolydūs kintamieji.

Stačiakampėje koordinatinių sistemoje vaizduojama Y reikšmių priklausomybė nuo kintamojo X . Ši priklausomybė gali būti išreikšta taip:

$$s(X) = E(Y | X). \quad (7)$$

Taip pat ši priklausomybė gali būti užrašoma kaip

$$y_i = s(x_i) + \varepsilon_i, E\varepsilon_i = 0, i = 1, 2, \dots, n. \quad (8)$$

Neapibrėžta glodi funkcija s , kuri turi būti įvertinta iš imties duomenų (x_i, y_i) , $i = 1, 2, \dots, n$, tai ji vadinama neparimetrine regresijos funkcija. Jai sudaryti imamas skaičius κ ir apibrėžiama tokia funkcija:

$$(x - \kappa)_+ = \begin{cases} 0, & \text{kai } x - \kappa < 0, \\ x - \kappa, & \text{kai } x - \kappa \geq 0. \end{cases} \quad (9)$$

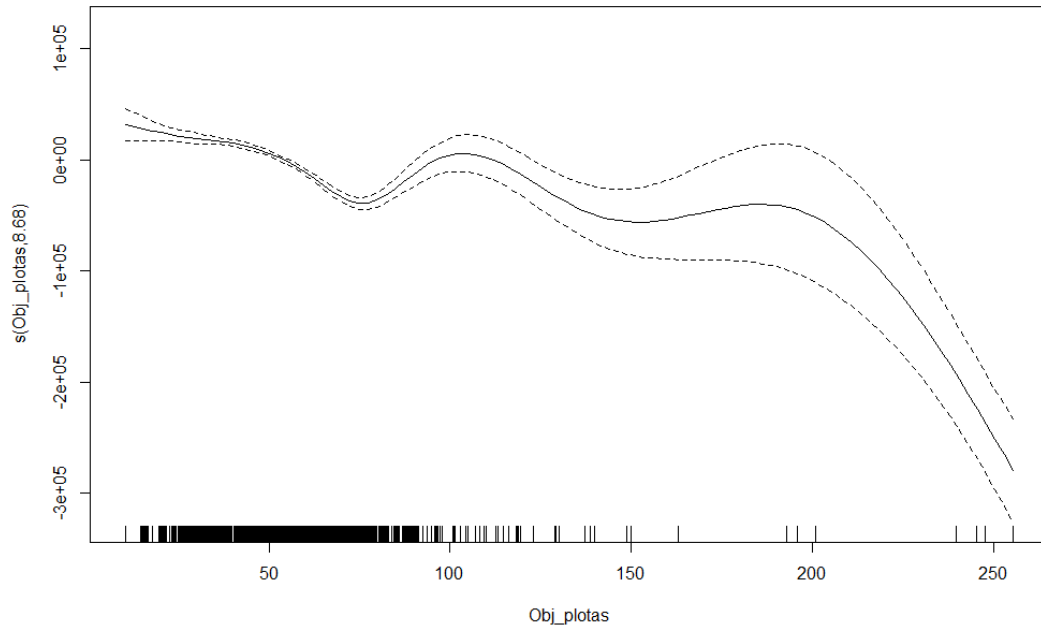
Skaičius κ vadinamas funkcijos $(x - \kappa)_+$ mazgu. Funkcija $(x - \kappa)_+$ yra vadinama tiesinio splaino bazine funkcija, o tokių funkcijų aibė vadinama tiesinio splaino baze. Bet kuri splaino bazinių funkcijų $1, x, (x - \kappa_1)_+, \dots, (x - \kappa_K)_+$ tiesinė kombinacija yra tiesinė funkcija su mazgais $\kappa_1, \kappa_2, \dots, \kappa_K$. Ši funkcija vadinama splainu.

Tokiu būdu gaunamas šiame darbe taikomas funkcijos $s(x)$ splaino modelis:

$$s(x) = \beta_0 + \beta_1 x + \sum_{k=1}^K \beta_k (x - \kappa_k)_+. \quad (10)$$

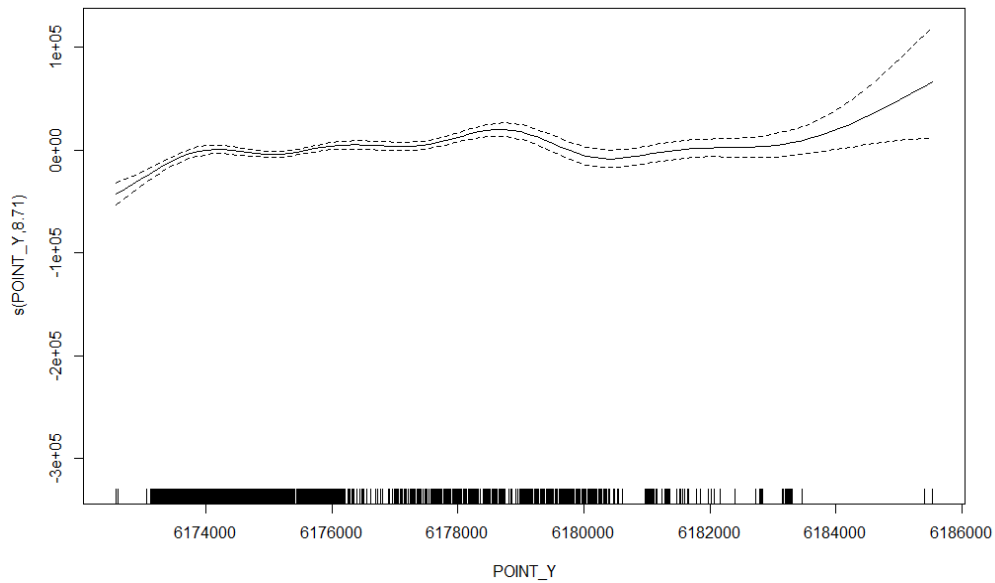
Remiantis kintamaisiais, iš kurių buvo sudarytas netiesinės regresijos modelis, sudaromas apibendrintasis adityvus modelis, pridėdamas kintamųjų objekto ploto ir geografinių koordinatinių glodžias funkcijas, bei priklausomam kintamajam pritaikyta logaritminė transformacija. Gautas modelis toliau žymimas Modeliu 5 ir šio modelio parametrų įverčiai pateikti 3 lentelėje.

Grafikuose (žr. 3–5 pav.) pavaizduoti įvertinti objekto ploto ir geografinių (X ir Y) koordinatinių kintamųjų glodžių funkcijų poveikiai modeliui ištiesine kreive su 95 % pasiklovimo intervalu, pažymėtu punktyrine linija.

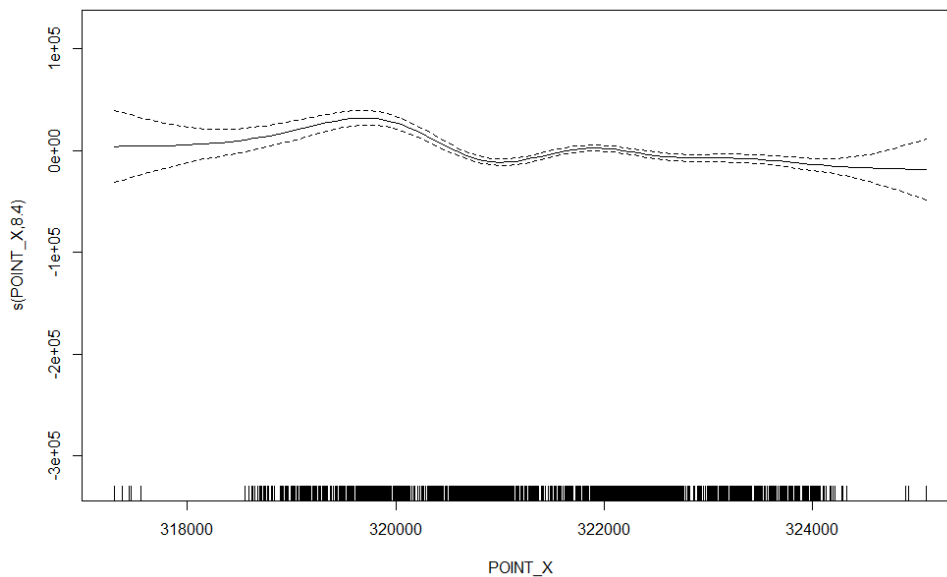


3 pav. Objekto ploto glodi funkcija

Pastebėtina, kad esant mažam sandorių kiekiui, šiuo atveju, kai objekto plotas yra daugiau nei 200 kv. m, tai glodžioji funkcija palaipsniui staigiai krinta.

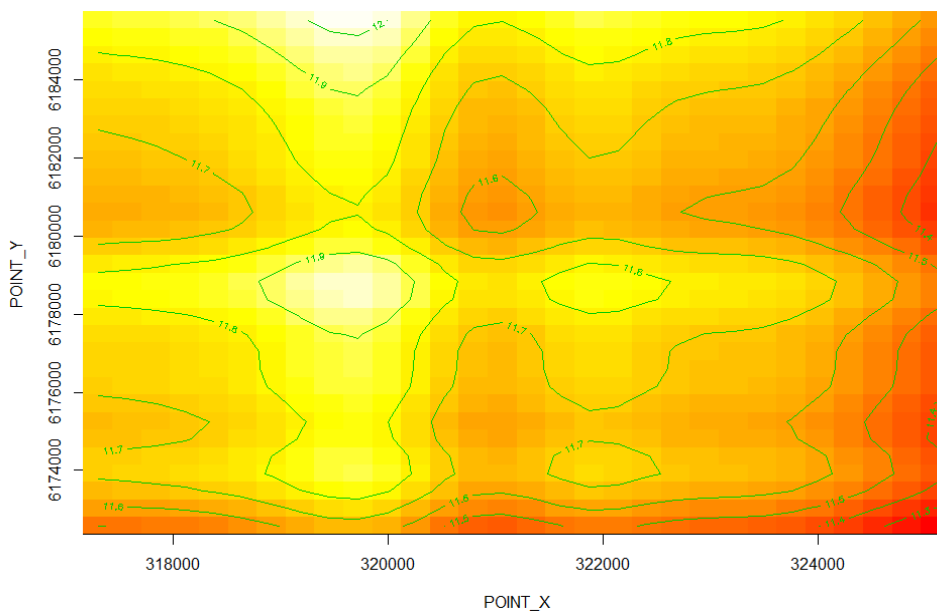


4 pav. Kintamojo Y koordinatės glodi funkcija



5 pav. Kintamojo X koordinatės glodi funkcija

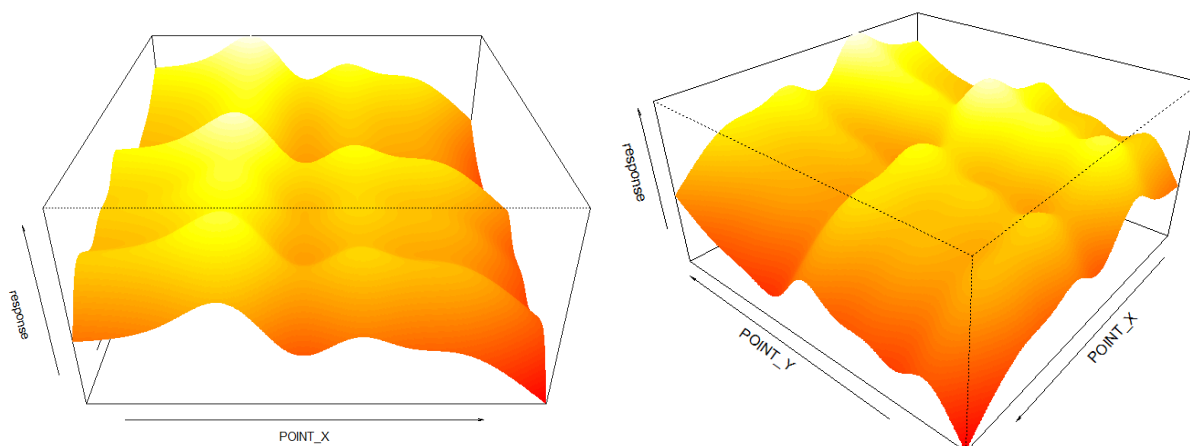
Kai į modelį yra įtrauktos geografinės koordinatės, rezultatus patogų vaizduoti pateikiant kontūrinį grafiką. Jame iš spalvų matyti, kaip priklausomojo kintamojo reikšmės (pardavimo kainos) priklauso nuo geografinių koordinatė, t. y. kokioje miesto dalyje butų kainos aukštos, o kurioje žemesnės. Šiuo atveju remiantis pateiktu kontūriniu grafiku (žr. 6 pav.), šviesiausia spalva vaizduojamos didžiausios butų kainos, o raudona – mažiausios butų kainos. Galima išvelgti tendenciją, jog tuo atveju, kai Klaipėdos miesto butai išsidėstę arčiau Baltijos jūros, tai ir prognozuotos butų kainos yra aukščiausios, o tolstant nuo jūros – žemiausios.



6 pav. 5 modelio kontūrinis grafikas

Prognozuotas Klaipėdos miesto butų kainas galima vaizduoti paviršiumi, nubrėžtu X ir Y stačiakampėje koordinatė sistema (žr. 7 pav.). 7 paveikslėlyje pateiktas 5 modelio kontūrinis grafikas, nubrėžiant X ir Y stačiakampėje koordinatė sistema. Į modelį įtrauktos geografinės koordinatės patikslina ne tik nekilnojamojo turto

objekto vietos įtaką, bet ir eliminavus iš vertinimo modelių zonas nebelieka šuolių pereinant nuo vienos zonos prie kitos.



7 pav. Prognozuotų butų kainų vaizdavimas X ir Y stačiakampėje koordinatinių sistemoje

4. Rezultatų palyginimas

Iš viso nagrinėti 5 regresijos modeliai Klaipėdos miesto butų kainoms vertinti. Gauti modeliai palyginami tarpusavyje, remiantis modelio tinkamumo rodikliais (vidutinės absoliutinės procentinės paklaidos rodikliu ir Akaikės informaciniu kriterijumi) (2 lentelė). Išrenkamas tinkamiausias modelis.

2 lentelės pirmame stulpelyje (Modelio nr.) pateikti nagrinėjamų modelių sutrumpinti žymėjimai, kur 1 modelis žymi netiesinės regresijos modelį, 2 modelis – tiesinę kintamųjų priklausomybę, išreikštą remiantis netiesinės regresijos modeliu (1 modeliu), 3 modelis – apibendrintąjį tiesinį regresijos modelį, taikant Bokso ir Kokso transformaciją priklausomam kintamajam, 4 modelis – apibendrintąjį tiesinį regresijos modelį, nurodžius logaritminę jungties funkciją, o 5 modelis – apibendrintąjį adityvų regresijos modelį (pusiau parametrinį regresijos modelį), pritaikius logaritmo transformaciją priklausomam kintamajam.

Modelio tinkamumui vertinti pasirinktas vidutinės absoliutinės procentinės paklaidos (ang. *mean absolute percentage error* – MAPE) rodiklis, kuris apskaičiuojamas taip:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\varepsilon_i|}{Y_i} \cdot 100\% . \tag{11}$$

Vidutinė absoliutinė procentinė paklaida nusako santykinę prognozavimo tikslumą ir, juo remiantis, galima palyginti skirtingų modelių prognozes. Iš gautų vidutinių absoliutinių procentinių paklaidų matyti, kad visais atvejais rodikliai yra panašūs, tačiau mažiausią reikšmę turi 5 modelis (pusiau parametrinis modelis), o didžiausią – 4 modelis. Kitas svarbus rodiklis modelio tinkamumui tikrinti – Akaikės informacinis kriterijus (*AIC*), kuris parodo sudaryto modelio kokybę, t. y. kuo šio kriterijaus reikšmė mažesnė, tuo modelis tinkamesnis. Remiantis sudarytais modeliais ir Akaikės informacinio kriterijaus reikšmėmis, galima lyginti tik tuos modelius tarpusavyje, kurių priklausomo kintamojo reikšmės yra panašios. Šiuo atveju 2 modelis ir 5 modelis lyginami tarpusavyje, nes šių modelių priklausomam kintamajam buvo pritaikyta logaritmo transformacija.

2 lentelė. Modelių statistinių rodiklių palyginimas

Modelio nr.	Vidutinė absoliutinė procentinė paklaida (MAPE)	Akaikės informacinis kriterijus (AIC)
1 modelis	20,38	93 601,30
2 modelis	19,09	233,61
3 modelis	19,46	18 340,39
4 modelis	20,79	106 029,32
5 modelis	18,45	-22,75

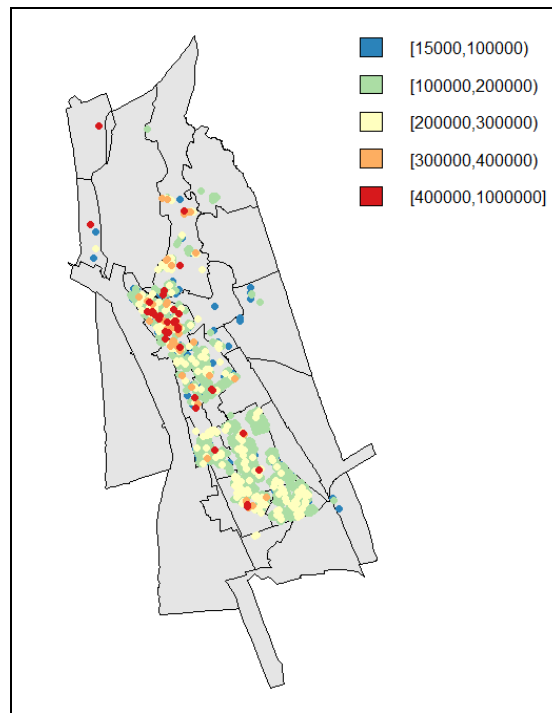
Palyginimui modelių parametrų įverčiai ir p reikšmės (pasirinktas reikšmingumo lygmuo 0,05) pateikiami 3 lentelėje. Pastebėtina, kad į 3 modelį ir 4 modelį įeina kintamųjų sąveikos, o 5 modelyje įtrauktos glodžios funkcijos (pažymėta pliuso „+“ ženklu). Lyginant 2 modelį, 3 modelį, 4 modelį ir 5 modelį tarpusavyje, gauta, jog šildymo kintamasis yra nereikšmingas. Tačiau visų modelių atvejais išskiriami tokie kintamieji kaip senumo grupė (namo statybos pabaigos metai), sienų medžiaga – plytos, 2 ir 3 kambarių butai, viršutinis aukštas ir objekto plotas. Vietos įtaka yra vienas iš pagrindinių rodiklių, turintis didžiausią svorį vertinant nekilnojamojo turto kainas. Atkreiptinas dėmesys, kad nagrinėjamame regresijos modelyje (5 modelis) neliko vietos įtakos (zonų) kintamųjų, tačiau Klaipėdos miesto butų vietos įtaką atspindi į modelį įtrauktos geografinės koordinatės, išreikštos per glodžias funkcijas.

3 lentelė. Modelių parametrų įverčių palyginimas

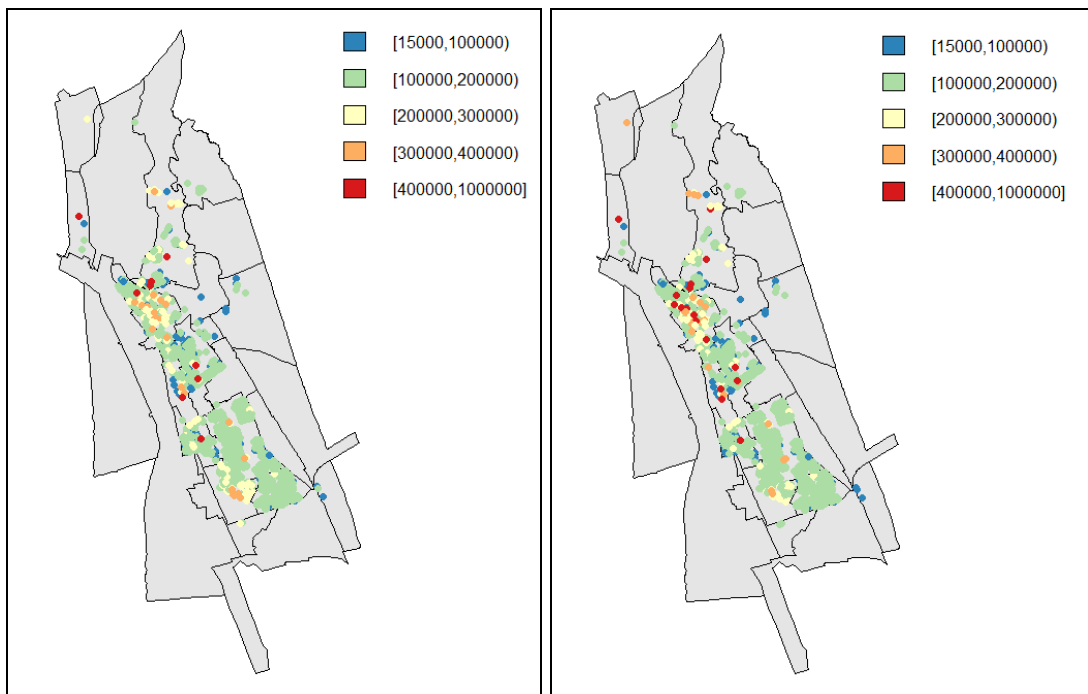
Kintamasis	1 modelio parametro įvertis	2 modelio parametro įvertis	3 modelio parametro įvertis	4 modelio parametro įvertis	5 modelio parametro įvertis
Laisvasis narys	-	-	32,41 [$<2e-16$]	11,006 [$<2e-16$]	1,297 [$<2e-16$]
Senumas2	0,86 [$<0,001$]	0,73 [$<2e-16$]	-2,36 [$<2e-16$]	-0,286 [$<2e-16$]	-0,363 [$<2e-16$]
Senumas3	0,88 [$<0,001$]	0,74 [$<2e-16$]	-2,61 [$<2e-16$]	-0,303 [$<2e-16$]	-0,424 [$<2e-16$]
Zonų grupė II	0,95 [$<0,001$]	0,95 [0,0002]	1,06 [7,9e-05]	0,276 [$<2e-16$]	-
Zonų grupė III	0,87 [$<0,001$]	0,85 [$<2e-16$]	-	-	-
Gelžbetonio plokštės	0,92 [$<0,001$]	-	2,27 [$<2e-16$]	0,334 [$<2e-16$]	-
Monolitinis gelžbetonis	1,10 [$<0,001$]	-	-	0,260 [0,0003]	-
Plytos	1,05 [$<0,001$]	1,06 [4,4e-11]	0,32 [0,0404]	0,091 [5,4e-09]	0,052 [2,1e-07]
Kamb2	0,95 [$<0,001$]	0,97 [0,0044]	0,81 [4,3e-12]	0,136 [$<2e-16$]	-0,039 [0,0003]
Kamb3	0,97 [$<0,001$]	0,95 [9,5e-06]	3,42 [5,2e-15]	0,399 [$<2e-16$]	-0,097 [7,3e-06]
Kamb4	0,89 [$<0,001$]	0,87 [1,7e-14]	4,95 [$<2e-16$]	0,722 [$<2e-16$]	-
Pirmas aukštas	0,93 [$<0,001$]	0,95 [8,1e-08]	-	-	-0,057 [1,2e-08]
Viršutinis aukštas	0,97 [$<0,001$]	0,96 [1,2e-05]	-0,26 [0,0005]	-0,033 [0,0008]	-0,043 [9,9e-06]
Nėra rūšio	0,99 [$<0,001$]	0,97 [0,0003]	-	-	-
Nėra dujų	1,17 [$<0,001$]	1,05 [0,0058]	-	-	0,049 [0,0045]
Krosninis	0,93 [$<0,001$]	-	-	-	-
Vietinis centrinis	1,06 [$<0,001$]	-	-	-	-
Objekto plotas	3538,87 [$<0,001$]	3944,19 [$<2e-16$]	0,17 [$<2e-16$]	0,019 [$<2e-16$]	0,211 [$<2e-16$]
Objekto plotas: Zonų grupė II	-	-	-0,03 [1,7e-08]	-0,005 [$<2e-16$]	-
Objekto plotas: Zonų grupė III	-	-	-0,02 [$<2e-16$]	-0,002 [$<2e-16$]	-
Objekto plotas: Pirmas aukštas	-	-	-0,01 [7,4e-08]	-0,001 [2,2e-08]	-
Objekto plotas: Nėra dujų	-	-	-	0,0003 [0,0409]	-
Objekto plotas: Kamb3	-	-	-0,04 [2,9e-07]	-0,003 [2,7e-06]	-
Objekto plotas: Kamb4	-	-	-0,07 [$<2e-16$]	-0,008 [$<2e-16$]	-
Objekto plotas: Gelžbetonio plokštės	-	-	-0,05 [$<2e-16$]	-0,006 [$<2e-16$]	-
Objekto plotas: Monolitinis gelžbetonis	-	-	-	-0,002 [0,0317]	-
s(Objekto plotas)	-	-	-	-	+ [$<2e-16$]
s(X koordinatė)	-	-	-	-	+ [$<2e-16$]
s(Y koordinatė)	-	-	-	-	+ [$<2e-16$]

Vienas iš būdų palyginti butų kainas (sandorio sumą) ir prognozuotas butų kainas yra pateikti jų reikšmės žemėlapyje. Pasinaudojant geografinėmis koordinatėmis, pavaizduojamos Klaipėdos miesto butų pirkimo ir pardavimo sandorių kainos 8 pav., o pasirinktų 2 modelio ir 5 modelio prognozuotos kainos – 9 pav. atitinkamai. Šių grafikų pateikimo tikslas – Klaipėdos miesto prognozuotų butų kainų pasiskirstymas būtų kuo panašesnis grafiniu pateikimu į realias butų kainas. Tačiau norint palyginti turimas kainas (sandorio sumas) su prognozuotomis kainomis, pirmiausia reikia kainas suskirstyti į atitinkamus intervalus. Parinktiems intervalams priskirtos atitinkamos spalvos – t. y. pigiausių

butų kainos vaizduojamos mėlyna spalva, pigesnių – žalsva, vidutinių (nuo 200 000 Lt iki 300 000 Lt) – gelsva, brangesnių – oranžine, o brangiausių – raudona.



8 pav. Butų kainų (sandorio sumos) pasiskirstymas Klaipėdos miesto savivaldybėje



9 pav. Prognozuotų butų kainų pasiskirstymas Klaipėdos miesto savivaldybėje (2 modelis ir 5 modelis)

Lyginant du prognozuotų kainų grafikus (žr. 9 pav.) su realiomis kainomis (žr. 8 pav.), geriausiai atspindi 5 modelio prognozuotos kainos, t. y. taikant apibendrintuosius adityvius modelius (pusiau parametrinius regresijos modelius). Šis modelis duoda tiksliausius rezultatus dėl įtrauktų į modelį glodžių funkcijų.

5. Išvados

Taikant siūlomą Tarptautinės turto vertintojų mokesčiams asociacijos nekilnojamojo turto vertinimo modelį, sudarytą taikant netiesinės regresijos modelį, modelis transformuojamas į tiesinį modelį, kai yra tik vienas kiekybinis kintamasis.

Į modelį įtrauktos geografinės koordinatės kaip kintamieji, taikant apibendrintuosius adityvius modelius, nusako tikslesnę nekilnojamojo turto objekto buvimo vietą ir tai įtakoja prognozuotas Klaipėdos miesto buto kainas. Tai būtų išvengiama skirtingų kainų lygio šuolio pereinant nuo vienos zonos prie kitos. Todėl ateityje sudarant vertinimo modelius Lietuvoje būtų galima tobulinti, atsižvelgiant į geografinių koordinatžių koreliaciją.

Remiantis sudarytų modelių atvejais, vienas iš tinkamiausiai įvertintų vertinimo modelių gautas taikant pusiau parametrinį regresijos modelį. Tačiau vertinimo modelio rezultatai gali priklausyti nuo nekilnojamojo turto požymių, duomenų stygiaus ar nekilnojamojo turto objekto buvimo vietos. Nagrinėjant Klaipėdos miesto butų kainas gautas tinkamas ir pakankamai adekvatus nekilnojamojo turto vertinimo modelis, kuris galėtų būti tiesiogiai naudojamas analizuojant ir prognozuojant Klaipėdos miesto butų kainas.

Literatūra

1. Aleknavičius, A. 2008: *Nekilnojamojo turto vertinimas*, Kaunas, Ardiva.
2. Clark, M. 2012: *Generalized Additive Models. Getting started with additive models in R*, University of Notre Dame: Center for Social Research, [interaktyvus], [žiūrėta 2014-02-22]. Prieiga per internetą: <http://www3.nd.edu/~mclark19/learn/GAMS.pdf>.
3. Crawley, M. J. 2005: *Statistics: An Introduction Using R*, UK: Wiley and Sons, Inc.
4. Crawley, M. J. 2007: *The R Book*, UK: Wiley and Sons, Inc.
5. Čekanavičius, V.; Murauskas, G. 2000, 2002, 2009: *Statistika ir jos taikymai, I, II, III*, Vilnius, TEV.
6. Dobson, A. J. 2002: *An Introduction to Generalized Linear Models*, Florida: Chapman & Hall/CRC.
7. Eckert, J. K. 1990: *Property Appraisal and Assessment Administration*, Chicago: The International Association of Assessing Officers.
8. García-Ponce, O. 2013: *GIS with R (or without ArcGIS)*, Spring Lab Workshop., New York: New York University.
9. Guszczka, J. 2010: *Geo-spatial Analysis with Generalized Additive Models*, Chicago, [interaktyvus], [žiūrėta 2014-02-22]. Prieiga per internetą: <http://www.casact.org/education/rpm/2010/handouts/PM1-Guszczka.pdf>.
10. Kasperavičius, R. 2012: *Nekilnojamojo turto administravimo sistema*, Vilnius, VĮ Registrų centras.
11. Kuhnert, P.; Venables, B. 2005: *An Introduction to R: Software for Statistical Modelling & Computing*. Australia: CSIRO Mathematical and Information Sciences.
12. Lietuvos Respublikos terminų bankas, [interaktyvus], [žiūrėta 2014-04-11]. Prieiga per internetą: http://terminai.vlkk.lt/pls/tb/tb_search.
13. Lietuvos Respublikos turto ir verslo vertinimo pagrindų įstatymas (Žin., 1999, Nr. 52-1672; 2011, Nr. 86-4139).
14. Lietuvos Žemės Ūkio Universitetas, Žemėtvarkos katedra. 2007: *Geografinių duomenų analizė*, Kaunas, [interaktyvus], [žiūrėta 2014-03-30]. Prieiga per internetą: <http://www.hidro.lzuu.lt/vuzf/metodiniai/Zemetvarka/GISaM.pdf>.
15. NCSS statistical software, [interaktyvus], [žiūrėta 2012-11-16]. Prieiga per internetą: www.ncss.com.
16. Nacionalinė žemės tarnyba prie Žemės ūkio ministerijos, VĮ Registrų centras. 2014: *Nekilnojamojo turto registre įregistruotų statinių apskaitos duomenys 2014 m. sausio 1 d.*, Vilnius.
17. Navickaitė, R. 2011: Baigiamasis bakalauro darbas. *Butų vertės prognozavimas, naudojant tiesinės ir netiesinės regresijos modelius*, Vilnius: VU.
18. Nekilnojamojo turto vertinimo taisyklės, patvirtintos Lietuvos Respublikos Vyriausybės 2005 m. rugsėjo 29 d. nutarimu Nr. 1049 (Žin., 2005, Nr. 117-4234; 2011, Nr. 28-1321).
19. Olsson, U. 2002: *Generalized Linear Models. An Applied Approach*, Sweden.
20. Ritz, Ch.; Streibig, J. C. 2008: *Nonlinear Regression with R*, New York: Springer.
21. Ruppert, D., Wand, M. P., Carroll, R. J. 2003: *Semiparametric Regression*, New York: Cambridge University Press.
22. The R Project for Statistical Computing, [interaktyvus], [žiūrėta 2012-11-16]. Prieiga per internetą: <http://www.r-project.org/>.
23. Turner, H. 2008: *Introduction to Generalized Linear Models*. UK: ESRC National Centre for Research Methods, UK and Department of Statistics University of Warwick.
24. VĮ Registrų centro internetinis tinklapis, [interaktyvus], [žiūrėta 2012-11-16]. Prieiga per internetą: www.registrucentras.lt.
25. VĮ Registrų centro žemės sklypų ir statinių masinio vertinimo dokumentai [interaktyvus], [interaktyvus], [žiūrėta 2012-11-16]. Prieiga per internetą: <http://www.registrucentras.lt/ntr/masvert/index.php>.

-
26. Wood, S. N. 2006: *Generalized Additive Models: an introduction with R*, USA: Chapman and Hall/CRC Press, [interaktyvus], [žiūrėta 2014-02-22]. Prieiga per internetą: <http://reseau-mexico.fr/sites/reseau-mexico.fr/files/igam.pdf>.

APPLICATION OF STATISTICAL METHODS TO THE ASSESSMENT OF PRICES OF KLAIPEDA CITY APARTMENTS

Roberta Navickaitė

Abstract. The paper analyses the use of a nonlinear regression model, generalised linear model and generalised additive model (semi-parametric regression model) for creating real estate valuation models. These models are applied to data on transactions in Klaipeda city apartments. The aim is to create real estate valuation regression models applying various statistical methods and to compare them with each other. The practical aspects of creating regression models are analysed and conclusions are presented in the paper.

Keywords: generalised additive model, generalised linear model, apartments, nonlinear regression model, semi-parametric regression model.