
IŠ DALIES EUROPOS SOCIALINIO FONDO LĖŠOMIS FINANSUOJAMAS PROJEKTAS "LIETUVOS GEOGRAFINĖS INFORMACIJOS
VALDYTOJŲ KVALIFIKACIJOS KĖLIMAS" (NR. BPD2004-ESF-2.2.0-02-05/0143)

Valstybės tarnautojų mokymo programa

GEOGRAFINĖS INFORMACIJOS INFRASTRUKTŪROS NUOTOLINIS MOKYMAS

Mokomoji knyga

ERDVINĖ ANALIZĖ IR MODELIAVIMAS

GII-07

Vilnius, 2008

Mokomoji knyga „Erdvinė analizė ir modeliavimas“ (GII-07)

Autoriai

1 skyrius -Ann Blyth, Dave Cake

2 skyrius- Ann Blyth, Ian Laing, Dave Cake

3 skyrius- Martin Andresen

4 skyrius- dr. Gennady Gienko and Michael Govorov

5 skyrius- Ian Laing, Dave Cake

Recenzavo ir redagavo

doc. dr. Gintautas Mozgeris (Lietuvos Žemės ūkio universitetas)

Recenzavo

doc. dr. Antanas Dumbrasuskas (Lietuvos Žemės ūkio universitetas)

Iš anglų kalbos vertė ir redagavo

UAB Astraneta

Valstybės tarnautojų nuotolinio mokymo programos „Geografinės informacijos infrastruktūros nuotolinis mokymas“, patvirtintos Valstybės tarnybos departamento prie LR Vidaus reikalų ministerijos direktoriaus įsakymu „Dėl valstybės tarnautojų mokymo programų tvirtinimo“ Nr. 27V-72 (2007-04-17), mokymo kursų medžiaga parengta vykdant iš dalies Europos Sąjungos lėšomis finansuojamą projektą „Lietuvos geografinės informacijos valdytojų kvalifikacijos kėlimas“ (Nr. BPD2004-ESF-2.2.0.-02-05/0143)

Projekto vykdytojas Nacionalinė žemės tarnyba prie Žemės ūkio ministerijos.

Projekto rangovas UAB HNIT-BALTIC, Vilniaus Gedimino technikos universitetas

© Autoriaus išimtinės turtinės teisės priklauso Nacionalinei žemės tarnybai prie Žemės ūkio ministerijos

© Autoriaus neturtinės teisės priklauso Malaspina University-College

Turinys

Kurso apžvalga.....	4
1. Duomenų tyrimas	6
1.1 Duomenų atributai	7
1.2 Vektorinių duomenų užklausos ir atranka	17
1.3 Rastrinių duomenų užklausos ir atranka	28
1.4 Duomenų sumavimas ir interpretavimas.....	32
2. Erdvinės analizės pagrindai.....	39
2.1 Vektorinės analizės metodai	40
2.2 Rastrinės analizės metodai.....	52
2.3 Duomenų apibendrinimas	70
3. Erdvinė statistika	79
3.1 Klasikinė koreliacija, autokoreliacija erdvėje ir vietos koeficientas.....	80
3.2 Kintamo plotinio vieneto problema ir ekologinė klaida	85
3.3 Erdvinio išsidėstymo analizė.....	87
3.4 Kraštų įtaka.....	95
3.5 Tankio skaičiavimas ir karštųjų zonų nustatymas	98
3.6 Vietinė erdvinė statistika	104
4. Geostatistika.....	111
4.1 Geostatistikos pagrindai.....	112
4.2 Geostatistinių metodų klasifikacija	118
4.3 Interpoliacija, vertinimas ir glodinimo metodai	120
4.4 Interpoliavimas krigingo metodu	129
4.5 Modelio patikra	154
4.6 Geostatistinių metodų palyginimas	164
5. Erdvinis modeliavimas.....	169
5.1 Erdvinių problemų modeliavimas	170
5.2 Erdvinių modelių klasifikavimas	173
5.3 Modelių tipai.....	175
5.4 Modelių kūrimo programa Model Builder	177
5.5 Modelių pavyzdžiai	179

Kurso apžvalga

GIS apibūdinama kaip priemonių rinkinys, skirtas rinkti, saugoti, analizuoti ir vaizduoti geografinius duomenis. Daugiausia dėmesio GIS skiriama užduotims, susijusioms su galimybe atvaizduoti ir aprašyti tikrojo pasaulio objektus skaitmeninėje terpėje. Šios srities temos susijusios su objektų abstrahavimu GIS duomenų bazėje taškais, linijomis ir poligonais, vietai žemės paviršiuje nurodyti skirtų koordinatinių sistemų ir kartografinių projekcijų supratimu bei naudojimu ir fiziniiais failų, kuriuose saugomi elementai erdvinio duomenų bazėje, formatais. Dauguma taikomųjų GIS programų funkcijų skirtos būtent minėtoms užduotims. Šios temos nagrinėjamos keliuose šios mokymo programos kursuose, pavyzdžiui:

GII-01	GIS pagrindai
GII-05	Geografinės DBVS
GII-06	Geodezija ir kartografija

Šios GIS funkcijos daugiausia susijusios su erdvinio objektų inventorizacija, o GIS analizės funkcijos skirtos padėti suprasti procesus ir dėsningumus, slypinčius už erdvinio duomenų bazėje atvaizduotų elementų. Būtent šios analizės galimybės ir skiria GIS nuo panašių sistemų, pavyzdžiui, kompiuterizuoto projektavimo (CAD) ir automatizuotos kartografijos. Erdvinė analizė gali padėti mokslininkams suprasti procesą arba objektų pasiskirstymą, organizacijai – priimti geresnius, gilesniu duomenų supratimu pagrįstus sprendimus.

Erdvinės analizės temų yra keliuose šios mokymo programos kursuose:

GII-01	GIS pagrindai
GII-04	Geografinės informacijos infrastruktūros taikymai
GII-07	Erdvinė analizė ir modeliavimas (šis kursas)

GII-01, įvadiniamame GIS kurse, bendrai apžvelgiama dauguma GIS temų, įskaitant ir erdvinę analizę. Todėl pastarojo kurso turinys iš dalies persidengs šio (GII-07) kurso turiniu, ypač su 2 dalimi. Bet šiame kurse minėtoji tema nagrinėjama išsamiau negu GII-01. GII-04 bus aptariamos konkrečios GIS taikymo sritys ir kaip šių temų dalis nagrinėjamos ypač specifinės analizės technikos, pavyzdžiui, tinklų analizė, reljefo modeliavimas ir hidrologinis modeliavimas. O šiame kurse daugiausia dėmesio skiriama pagrindinėms bendroms analizės technikoms. Taigi, nors GII-04 ir GII-07 kursuose gali kartotis tos pačios pagrindinės sąvokos, jų turinys persidengia nedaug.

Šio kurso tikslas – supažindinti su erdvinio duomenų analizės technikomis. Tipinėje GIS aplikacijoje yra įvairiausių analizės priemonių – nuo paprastų operacijų, tokių kaip atstumo matavimai, iki sudėtingų modelių skirtų numatyti rezultatus erdvėje pagal esamas arba numatomas pradines sąlygas. Kadangi erdvinės analizės technikų labai daug, daugumą jų galima išnagrinėti tik paviršutiniškai.

1 kurso dalyje apžvelgiamos pačios pagrindinės analizės priemonės, skirtos turimiems duomenims vizualizuoti arba nagrinėti. Aptarsime paprastas užklauso ir atrankos procedūras ir elementarius turimų duomenų sumavimo būdus – pagrindines statistikas, santraukas ir teminius žemėlapius.

2 dalyje nagrinėjama tai, ką daugelis pavadintų pagrindinėmis analizės priemonės. Tai geografinio apdorojimo įrankiai ir technikos, nuolat naudojami kiekvienoje GIS darbo vietoje – perdangos, buferių kūrimo ir karpymo operacijos. Šioje dalyje bus nagrinėjamos ir rastrinės, ir vektorinės analizės priemonės.

3 dalyje bus aptarti erdviniams duomenims taikomi statistiniai metodai. Jų pagrindas yra klasikinė statistika, bet jie įvertina erdvinių duomenų savybes ir apribojimus. Šios priemonės ypač naudingos tada, kai reikia suprasti ir aprašyti didelius erdvinių duomenų kiekius. Šioje dalyje apžvelgsime autokoreliaciją erdvėje, išsidėstymo analizę, tankio skaičiavimo metodus ir panašias temas.

4 dalyje nagrinėjamas tolydžių geografinių reiškinių tyrimo procesas. Šioje geostatistikos dalyje nagrinėjami interpoliavimo, glodinimo ir stebėjimų numatymo procesai bei kelios technikos, tokios kaip atvirkščiai proporcingas atstumas (IDW), aproksimavimas polinomais bei splineais ir krigingas.

5 dalyje nagrinėjami erdvinėje analizėje naudojami modeliai ir jų taikymas. Šioje dalyje aptarsime daug įvairių modelių tipų ir išsamiai išnagrinėsime kelis modelius

1. Duomenų tyrimas

Geografinės informacinės sistemos (GIS) unikalios tuo, kad jos gali susieti duomenis su jų vieta erdvėje ir užklausti šių duomenų bei juos sumuoti pagal konkrečius analizės poreikius. Funkciniu požiūriu GIS yra sudėtingas duomenų bazės ataskaitų sudarymo įrankis. Šios ataskaitos gali būti viso duomenų rinkinio (arba lentelės) arba duomenų rinkinio dalies (pvz., iš užklaustos arba duomenų sumavimo rezultatų). Duomenų ataskaitos gali būti suvestinių lentelių, diagramų arba žemėlapių formos. Šioje kurso dalyje mokysitės, kokie yra duomenų tipai, kaip atrinkti duomenų bazėje saugomą informaciją ir kaip jos užklausti. Be to, išsamiai aptarsime galimus ataskaitų tipus. Duomenų tyrimas šioje dalyje suskirstytas į keturias temas:

- Duomenų atributai
- Vektorinių duomenų užklaustos ir atranka
- Rastrinių duomenų užklaustos ir atranka
- Duomenų sumavimas ir interpretavimas

Daugelyje potemių aptariamos specifinės techninės sąvokos. Kad šias sąvokas suprasti būtų lengviau, pateikiami ArcMap (9.2 versijos) naudotojo sąsaja sukurti pavyzdžiai. Nors šie pavyzdžiai sukurti naudojant ArcMap, iš principo jie tinka ir kitoms GIS naudotojo sąsajoms.

1.1 Duomenų atributai

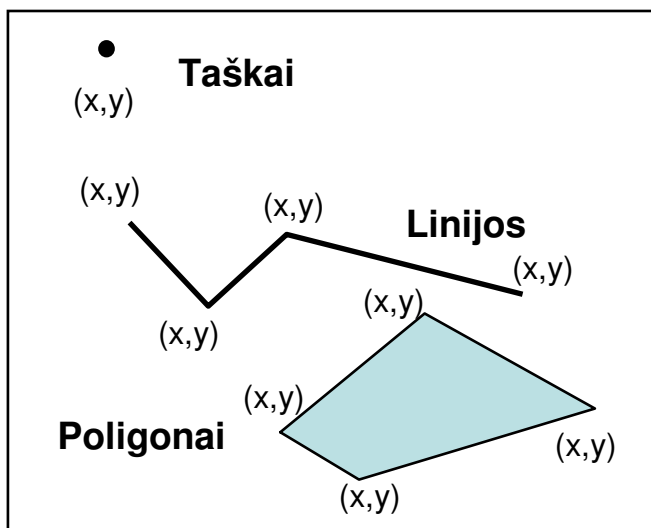
1.1.1 Vieta (pvz., X, Y ir Z atributai)

Pagrindinė erdvinės analizės sąvoka yra vieta – kur žemės paviršiuje yra tam tikras objektas ar objektų grupė. Daugiausia erdvinės analizės atliekama su dvimatėje erdvėje esančiais objektais. Sudėtingesnę analizę galima atlikti su trimačiais duomenimis (pvz., objektais po žemės paviršiumi arba virš jo). Kartais GIS duomenų rinkiniuose atsižvelgiama ir į ketvirtą matavimą – laiką. Daugialaikio duomenų rinkinio pavyzdys galėtų būti žemėnaudos duomenys, kuriuose saugomi kelių metų žemėnaudos atributai, kad būtų galima analizuoti žemėnaudos kaitą laikui bėgant.

Duomenys saugomi keturiais pagrindiniais formatais, iš kurių trys – vektorinis, rastrinis ir netaisyklingųjų trikampių tinklo (NTT) – skirti ir erdvinei informacijai saugoti. Be šių trijų, galima naudoti ir lentelių pavidalo duomenis, kuriuos vėliau galima susieti su erdvinio duomenų rinkiniais.

Vektoriniai duomenys

Vektoriniai duomenys konstruojami iš taškų. Šių taškų vietai nusakyti naudojama iki trijų koordinačių: x , y ir z . Taško vietai nusakyti reikia bent x ir y koordinačių poros. Z koordinatė gali reikšti papildomą su tašku susijusį dydį, pavyzdžiui, altitudę. Pačių koordinačių reikšmės priklauso nuo duomenų koordinačių sistemos (pvz., ilgumos ir platumos koordinatės arba Universali skersinė Merkatoriaus (UTM) sistema). Altitudės reikšmės (įprastinę z koordinatę) galima išreikšti įvairiais matavimo vienetais (metrais, pėdomis ir kt.). Vektoriniame duomenų modelyje taškiniai, linijiniai ir poligoniniai (plotiniai) erdviniai objektai sudaromi iš taškų (su jų x ir y koordinatėmis) (1.1 pav.). Taškai yra bemačiai, ir turi tik vietos savybę. Linijos yra vienmatės, ir turi ilgio savybę. Paprasčiausia linija nurodoma dviem taškais – pradiniu tašku ir galiniu tašku. Sudėtingesnių linijų forma nurodoma neribotu kiekiu taškų, esančių tarp pradinio ir galinio taško – kreivę galima paglondinti, padidinus šių taškų skaičių. Poligonai (plotiniai objektai) – tai uždaros linijos; jie yra dvimačiai. Šie elementai turi ploto ir perimetro savybes. Plotas gali būti atskiras arba jungtis su kitais elementais bendromis ribomis. Erdvinius ryšius tarp besijungiančių arba gretimų vektorinių objektų nurodo topologija. Ji aprašo, pavyzdžiui, kurios poligonų ribos yra bendros ir kaip linijos jungiasi viena su kita (pvz., kelių sankirtas).



1.1 pav. Vektoriniai duomenys: taškai, linijos ir poligonai

Taškų, linijų ir poligonų atributai priskiriami konkrečiam elementui ir saugomi arba kartu su erdvinio duomenų rinkiniu (paprasta duomenų struktūra), arba atskirose atributų lentelėse, kurios unikaliais identifikatoriais susiejamos su erdviniais elementais.

Įprastinių vektorinių modelių vaizduojamų duomenų pavyzdžiai yra diskretūs objektai (kuriuos galima išskirti kaip atskirus) – mėginių ėmimo vietos (paprastai saugomos kaip taškai), keliai (paprastai saugomi kaip linijos) ir sklypų ribos (sklypai paprastai saugomi kaip poligonai).

Vektoriniai duomenys ir vektorinis duomenų modelis turi savų pranašumų¹:

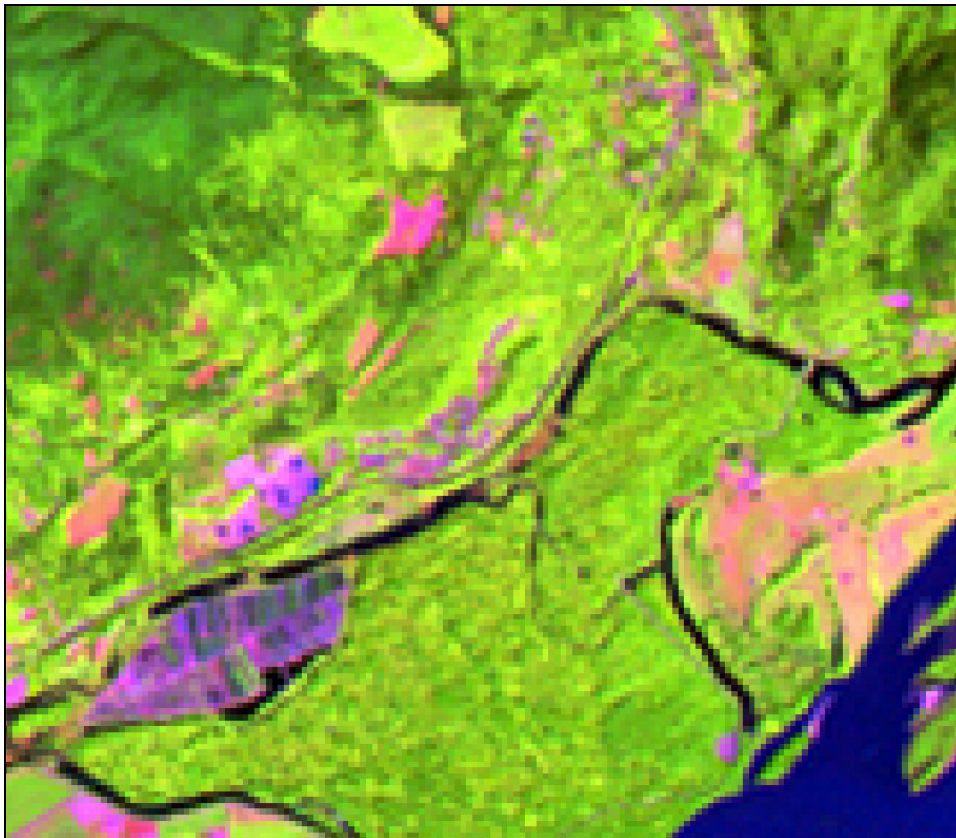
- galima efektyviai saugoti labai tikslius duomenis;
- lyginant su rastriniu formatu, tiems patiems duomenims saugoti reikia tik 10 proc. vietos;
- kai kurių tipų topologinę analizę geriau atlikti arba įmanoma atlikti tik su vektoriniais duomenimis;
- didesnis tikslumas ir preciziškumas;
- platesnės atributų duomenų saugojimo ir apdorojimo galimybės.

Rastriniai duomenys

Rastrinis duomenų modelis duomenis vaizduoja į vienodo dydžio gardelės padalintu paviršiumi. Paprastai rastras yra tinklelis, turintis pradžią (dažniausiai kairysis viršutinis kampas), ir kiekvienos gardelės vieta nurodoma poslinkiu nuo šio pradžios taško. Duomenų tinklelis paprastai stačiakampis, jo dydis nurodomas eilučių ir stulpelių skaičiumi; tinklelio aprėptis skaičiuojama dauginant tinklelio dydį (eilučių ir stulpelių skaičių) iš gardelės dydžio (metrinėje sistemoje). Nors rastrai gali būti įvairios formos (pvz., trikampiai, šešiakampiai), paprastai gardelės yra stačiakampės, dažniausiai – kvadratinės. Kiekvienoje gardelėje saugomas su atitinkama žemės paviršiaus dalimi susijęs atributas arba reikšmė. Rastrinis duomenų modelis idealiai tinka saugoti informaciją, susijusią su

¹ Kurie taip pat yra ir rastrinių duomenų trūkumai

tolydžiais objektais (objektais, kurie erdvėje nėra diskretūs), pavyzdžiui, apie temperatūrą arba altitudę. Vaizdai – pavyzdžiui, aeronuotraukos ir palydoviniai vaizdai – taip pat saugomi rastriniu formatu (1.2 pav.). Duomenys rašomi į paprastą duomenų struktūrą, kurią sudaro su fiksuotomis gardelės vietomis susieti eilutės ir stulpeliai.



1.2 pav. Rastrinės struktūros pavyzdys: palydovinis vaizdas

Kiekviena rastro gardelė turi aukštį ir plotį, kuris gali būti keli centimetrai, metrai ar net kilometras. Gardelių matmenys paprastai priklauso nuo duomenų skiriamosios gebos. Gardelės dydis lemia, ar duomenys bus atvaizduoti tiksliau, ar grubiau – paprastai kuo sudėtingesni duomenys, tuo daugiau gardelių reikia jiems tiksliai atvaizduoti. Atrodytų, kad tikslumą padidinti labai paprasta – reikia visiems duomenims nustatyti labai mažą gardelės dydį; bet kita vertus, kuo mažesnis gardelės dydis, tuo didesnis failas, kuriame duomenys saugomi, ir tuo ilgiau trunka juos atvaizduoti ekrane bei apdoroti. Paprastai gardelių dydis yra kompromisas tarp duomenų prieinamumo, tikslumo, užimamos vietos saugykloje ir vaizdavimo bei apdorojimo spartos. Gardelės dydis dar vadinamas duomenų skiriamąja geba: jei gardelė yra 25 × 25 metrai, duomenų rinkinio skiriamoji geba yra 25 metrai. Kuo mažesnė gardelė, tuo didesnė skiriamoji geba, ir tuo smulkesnės detales galima pavaizduoti.

Su kiekviena gardele susiejama vaizduojamą atributą atitinkanti reikšmė (pvz., su kiekviena kritulių duomenų rinkinio gardele bus susietas lietaus kiekis, o žemėnaudos gardelės turės žemėnaudos tipą nurodantį atributą – miesto, žemės ūkio, miško ir pan.). Gardelių reikšmės gali būti teigiami ir neigiami sveikieji arba slankaus kablelio skaičiai. Kartais gardelėms priskiriama reikšmė „NODATA“, nurodanti, kad duomenų nėra (1.3 pav.).



Zone with value 1	Zona, kurios reikšmė 1
Zone with value 2	Zona, kurios reikšmė 2
Zone with value 3	Zona, kurios reikšmė 3
Zone with value 4	Zona, kurios reikšmė 4
Zone with value 5	Zona, kurios reikšmė 5
NO DATA	DUOMENŲ NĖRA

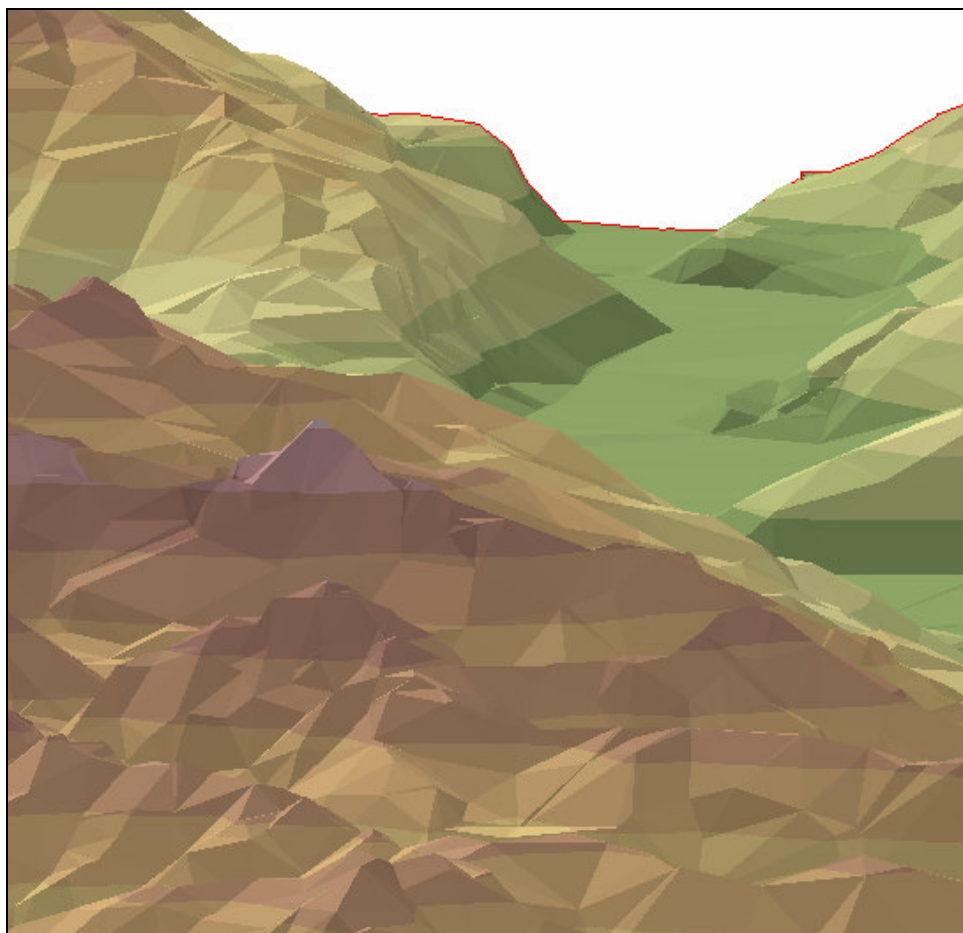
1.3 pav. Rastrinių duomenų pavyzdys

Rastriniai duomenys ir rastrinių duomenų modelis taip pat turi daug pranašumų, lyginant su vektoriniais duomenimis:

- duomenų modelis paprastas;
- daug erdvinės analizės funkcijų paprastesnės ir spartesnės;
- patogus didelio kintamumo erdvėje duomenims;
- suspaudus patogus mažo kintamumo erdvėje duomenims;
- lengva integruoti su palydoviniais ir kitais nuotolinių tyrimų duomenimis.

Netaisyklingųjų trikampių tinklas

Netaisyklingųjų trikampių tinklo (NTT) duomenų modelis idealiai tinka vaizduoti paviršių, pavyzdžiui, reljefą. Duomenys vaizduojami nepersidengiančių trikampių, nubrėžtų tarp nevienodais tarpais išdėstytų taškų, rinkiniu (1.4 pav.). Reljefo modelyje kiekvienas trikampis atitinka vienodo nuolydžio (gradiento) plotą. Kadangi NTT galima vaizduoti netaisyklingai išdėstytus duomenis, jais daug tiksliau negu rastrais galima modeliuoti tam tikrose vietose staigiai kintančius paviršius – kur duomenys kinta labiau, galima pridėti daugiau taškų, kur kintamumas mažesnis, užtenka mažiau taškų.



1.4 pav. Vietovės paviršiaus NTT pavyzdys

Lentelių duomenys

Lentelių duomenys naudojami atributams – erdvinių duomenų aprašomajai informacijai saugoti. Paprastai informacija saugoma duomenų bazėje eilučių ir stulpelių (laukų) pavidalu. Atributai gali būti toje pačioje lentelėje kaip ir erdvinė informacija, arba juos galima rašyti į atskiras lenteles ir unikaliais identifikatoriais susieti šias lenteles su erdvinių duomenų rinkiniu. GIS programinės įrangos paketai lentelių pavidalo duomenis gali saugoti įvairiais formatais – kaip tekstinius, *dBASE*, *Microsoft Excel* ar *Microsoft Access* failus. Be to, galima importuoti įvairių kitų duomenų bazių valdymo paketų (pvz., *Oracle*) išsaugotas lenteles.

1.1.2 Atributų tipai

Kaip jau aptarėme anksčiau, atributų duomenimis aprašomos erdvinių objektų savybės. Tačiau atributų duomenis galima klasifikuoti pagal duomenų tipą. Duomenų tipai įvairiose GIS programinės įrangos paketuose skiriasi, bet įprastiniai duomenų tipai yra simboliai (*character*), sveikieji skaičiai (*integer*), slankaus kablelio skaičiai (*float*), dešimtainiai (*decimal*), viengubo (*single*) ir dvigubo (*double*) tikslumo skaičiai bei eilutės (*string*). Įprastinių duomenų tipai, palaikomi GIS, pavaizduoti 1.1 lentelėje.

1.1 lentelė. Įprastiniai duomenų tipai*

Pavadinimas	Lauko intervalas / ilgis	Naudojimas
Tekstas (<i>Text</i>)	Iki 64 000 ženklų	Tinka pavadinimams, klasėms arba kitiems tekstiniams atributams
Data (<i>Date</i>)	mm/dd/yyyy hh:mm:ss A/PM	Data ir laikas
Trumpasis sveikasis skaičius (<i>Short integer</i>)	nuo -32 768 iki 32 767	Skaitiniai netrupmeniniai atributai (t. y. sveikieji skaičiai) nurodytame intervale. Šio tipo laukai tinka koduotoms reikšmėms
Ilgasis sveikasis skaičius (<i>Long integer</i>)	nuo -2 147 483 648 iki 2 147 483 647	Skaitiniai netrupmeniniai atributai nurodytame intervale (šio tipo laukuose galima saugoti ir dešimtaines reikšmes).
Viengubo tikslumo slankaus kablelio skaičiai (<i>Single precision floating point number</i>)	Maždaug nuo -3,4E38 iki 1,2E38	Skaitiniai trupmeniniai atributai nurodytame intervale (šio tipo laukuose galima saugoti ir dešimtaines reikšmes).
Dvigubo tikslumo slankaus kablelio skaičiai (<i>Double precision floating point number</i>)	Maždaug nuo -2,2E308 iki 1,8E308	Skaitiniai trupmeniniai atributai nurodytame intervale (šio tipo laukuose galima saugoti ir dešimtaines reikšmes).
BLOB	Skirtingi	Vaizdai ir kiti daugialypės terpės failai
GUID	36 ženklai	Globalūs identifikatoriai specifinėms aplikacijoms

*Laukų tipai ir reikšmių intervalai paimti iš ArcMap, bet pateikti pavyzdžiai tiks ir kitiems GIS programinės įrangos paketams.

Skaitinius laukus galima saugoti įvairiais skaitiniais duomenų tipais, pavyzdžiui, kaip trumpus sveikuosius skaičius arba ilgus sveikuosius skaičius ar kaip viengubo arba dvigubo tikslumo slankaus kablelio skaičius. Šie tipai skiriasi dydžiu ir skaitmeninių verčių saugojimo būdu (žr. 1.1 lentelę). Tekstiniuose laukuose saugomi tekstiniai simboliai (pvz., augalų pavadinimai). Datos duomenų tipo reikšmės gali būti data, laikas arba data ir laikas. Vaizdai, programų tekstai ir įvairūs daugialypės terpės failai saugomi kaip dideli dvejetainiai objektai (BLOB duomenų tipas – *Binary Large Object*). GlobalID ir GUID duomenų tipai yra registro stiliaus eilutės, unikalios identifikuojančios objektą arba lentelės eilutę vienoje ar keliose geoduomenų bazėse.

Atributinius duomenis apibūdina ir matavimo skalė. Šiuo atveju atributai gali būti kategorinio (pvz., nominaliniai (*nominal*) arba ranginiai (*ordinal*) duomenų tipai) arba skaitinio (pvz., intervaliniai (*interval*), santykiniai (*ratio*) arba cikliniai (*cyclic*) duomenys) tipo. Toliau pateikiame šių tipų apibrėžimus.

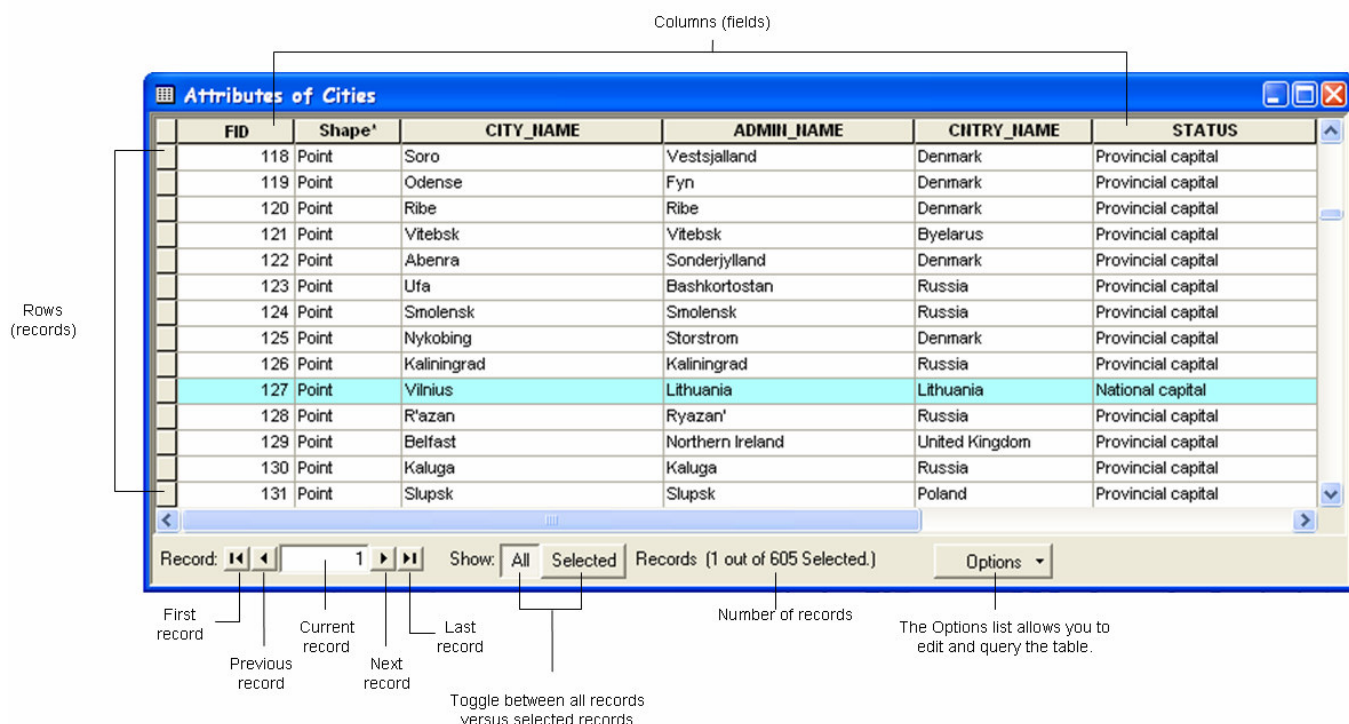
- **Nominaliniai** (vardiniai) duomenys aprašo įvairias duomenų kategorijas (pvz., žemėnaudos tipus arba augmeniją). Nominaliniai duomenys gali būti ir skaitiniai (pvz., telefonų numeriai arba žemėnaudos klasių skaitinės reikšmės), bet klasių reikšmės nesudaro jokios hierarchijos.

- **Ranginiai** duomenys reiškia, kad tarp klasių yra hierarchija – pavyzdžiui, eismo intensyvumą galima suskirstyti į aukšto, vidutinio ir žemo intensyvumo klases. Bet nors ir hierarchiniai, ranginiai duomenys yra kokybinio pobūdžio (t. y. jiems nepriskirta skaitinė reikšmė).
- **Intervaliniai** duomenys yra kiekybiniai, t. y. žinomos jų išmatuotos reikšmės arba intervalai tarp reikšmių. Intervalinių duomenų pavyzdžiai yra aukščio duomenys, kritulių lygis, temperatūra.
- **Santykiniai** duomenys taip pat yra kiekybinio pobūdžio. Jie panašūs į intervalinius, bet atskaitomi nuo prasmę turinčios (absoliučios) nulinės reikšmės. Santykinių duomenų pavyzdys yra tankio reikšmės, nes tankis gali kisti nuo 0 iki begalybės.
- **Cikliniai** duomenys yra kryptį arba ciklinius reiškinius vaizduojančių atributų reikšmės. Kartais du taškai matavimo skalėje gali sutapti. Pavyzdžiui, krypties duomenų atveju 0° ir 360° kryptis yra ta pati.

Duomenų tipas, kuriuo saugoma su tam tikru atributu susijusi informacija, priklauso nuo šiems duomenims atvaizduoti reikalingos matavimo skalės. Pavyzdžiui, nominaliniai arba ranginiai duomenys bus saugomi kaip tekstas, o intervaliniai, santykiniai arba cikliniai – kaip sveikieji arba dešimtainiai (slankaus kablelio) skaičiai. Kaip jau minėjome, nominalinius duomenis galima vaizduoti skaitinėmis reikšmėmis, paprastai sveikųjų skaičių tipo. Tačiau šiuo atveju skaičiai yra kodai, nurodantys peržiūros lentelėje (*look-up table*) saugomas tikrąsias reikšmes.

1.1.3 Atributų lentelės ir duomenų struktūros

Atributai saugomi atributų lentelėse. Paprastai vienoje lentelėje saugomi vienos panašaus pobūdžio objektų grupės atributai. Atributų lentelės sąranga vadinama jos duomenų struktūra. Atributų lentelės struktūruojamos į eilutes ir stulpelius. Stulpeliai atitinka lauką (atributo reikšmę) arba tam tikrą tipą. Pavyzdžiui, vandens mėginių ėmimo vietų duomenų rinkinys gali turėti vandens temperatūrą, pH reikšmes arba nuosėdų kiekį nurodančią atributų lentelę. Šios lentelės stulpeliai arba laukai gali būti identifikatorius, mažiausia temperatūra, didžiausia temperatūra, žiemos pH, vasaros pH ir kt. Lentelės eilutės atitinka individualias kiekvienos vietos (objekto) savybes (atributus). Duomenų bazėse eilutės taip pat vadinamos įrašais (pvz., kiekvienos vietos identifikatoriai ir temperatūros bei pH reikšmės). Lentelės stulpelių skaičius konkretus, o eilučių gali būti kiek nori. 1.5 pav. pavaizduota, kaip galima atributų lentelę parodyti ekrane. Pateiktas pavyzdys sukurtas naudojant *ArcMap*, bet sąvokos (pvz., eilutės, stulpeliai ir įrašai) tos pačios ir kitose GIS naudotojo sąsajose.



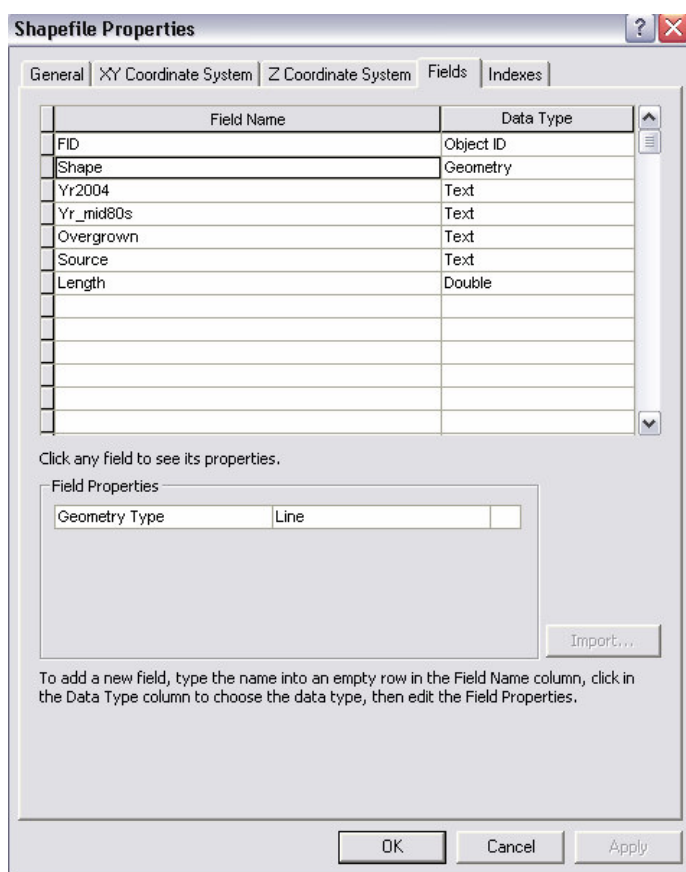
Columns (fields)	Stulpeliai (laukai)
Rows (records)	Eilutės (įrašai)
First record	Pirmasis įrašas
Previous record	Ankstesnis įrašas
Current record	Einamasis įrašas
Next record	Kitas įrašas
Last record	Paskutinis įrašas
Toggle between all records versus selected records	Rodyti visus arba tik atrinktus įrašus
Number of records	Įrašų skaičius
The options list allows you to edit and query the table	Lentelei redaguoti ir užklauskoms atlikti skirtas parinkčių sąrašas

1.5 pav. Atributų lentelės elementai

Per GIS naudotojo sąsają šiomis ar duomenų naršymo funkcijomis atributų lentelės galima peržiūrėti, keisti arba atlikti jų duomenų užklauskas. Be to, su konkrečiu elementu susijusius lentelių duomenis galima pasiekti iš žemėlapių, pažymėjus (paprastai spustelėjus pele) šį elementą (pvz., tašką, liniją arba poligoną). Pavyzdžiui, *ArcMap* programoje šioms duomenų užklauskoms skirtas identifikavimo įrankis. Kad įjungtumėte identifikavimo (*Identify*) funkciją, kairiuoju pelės klavišu spustelėkite identifikavimo mygtuką. Žemėlapių lange kairiuoju pelės klavišu spustelėkite dominantį elementą, ir ekrane bus parodyta tos vietos žemėlapių sluoksnių duomenų lentelė. Galima pasirinkti, kurių sluoksnių duomenis rodyti – viršutiniojo (*Top Layer*), matomų (*Visible Layers*), pažymimų (*Selectable Layers*) ar visų (*All Layers*), arba galima iš sąrašo pasirinkti konkretų sluoksnį.

1.1.4 Laukų savybės

Kuriant naują lentelę arba elementų klasę, savybių dialogo lange galima nurodyti, kiek laukų bus atributų lentelėje. 1.6 pav. pavaizduota, kaip šio tipo dialogas atrodo *ArcMap* naudotojo sąsajoje. Taip pat galima nurodyti laukų nuostatas, pavyzdžiui, lauko tipą ir didžiausią lauke saugomų duomenų dydį.



1.6 pav. SHAPE failo savybių dialogo langas

Galima nustatyti su esamu sluoksniu susietų laukų rodymo savybes. Paprastai galima: nurodyti, kurie laukai bus rodomi ekrane (t. y. įjungti ir išjungti laukų rodymą); priskirti laukui alternatyvų vardą, suprantamiau apibūdinantį lauko turinį; nurodyti skaitinių duomenų formatą (pvz., kiek skaičių po kablelio rodyti); nustatyti pagrindinį rodomą lauką.

1.1.5 Atributų jungimas ir siejimas

Atributus galima saugoti paprastuose failuose vienoje didelėje lentelėje (pvz., elementų atributų lentelėje) arba reliacinėje duomenų bazėje. Reliacinėse duomenų bazėse duomenys saugomi skirtingose lentelėse. Šias lenteles galima susieti tarpusavyje unikaliais identifikatoriais arba raktais. Reliacinės duomenų struktūros sumažina tų pačių duomenų kartojimo duomenų bazėje laipsnį. Be to, kadangi kiekviena lentelė saugoma atskirai, ją galima atskirai paruošti, prižiūrėti ir keisti. Reliacinės duomenų bazės dar patogios tuo, kad padidina ir duomenų valdymo, ir duomenų apdorojimo našumą – lenteles susieti tarpusavyje reikia tik atliekant tam tikrą užklausą arba analizę. Be šių pranašumų, reliacinės duomenų struktūros gali palengvinti duomenų mainus – jei keičiasi tik atributai, galima išplatinti tik naująją atributų lentelę, o erdviųjų duomenų rinkinio pakartotinai platinti nereikia.

Ryšiai tarp reliacinės duomenų bazės lentelių įrašų gali būti keturių tipų:

- **vienas su vienu** – vienas pirmosios lentelės įrašas susijęs su vienu (ir tik vienu) antrosios lentelės įrašu. Pavyzdžiui, į susietą lentelę galima įrašyti vandens kokybės

tikrinimo stočių atributų duomenis. Įrašų atributų lentelėje bus tiek, kiek yra stočių, o ryšys sudaromas stoties identifikatoriaus atributu, kuris yra bendras abiemis lentelėms.

- **vienas su daug** – vienas vienos lentelės įrašas susijęs su daug kitos lentelės įrašų. Pavyzdžiui, vienoje gatvėje gali būti keli namai. Gatvę vaizduojantis erdvinis elementas turės unikalų identifikatorių (pvz., GATVES_ID), ir namų įrašai šiuo identifikatoriumi bus susieti su šia gatve.
- **daug su vienu** daug vienos lentelės įrašų gali būti susiję su vienu kitos lentelės įrašu. Pavyzdžiui, žemėnaudos duomenų rinkinyje gali būti šimtai žemėnaudos poligonų, o atskiroje peržiūros lentelėje saugomų galimų žemėnaudos klasių – dvidešimt. Erdviniams duomenims priskiriamas tik poligono žemėnaudą atitinkantis skaitinis kodas nuo vieno iki dvidešimties. Šiuo kodu galima susieti abi minėtas lenteles. Šiuo atveju naudoti „daug su vienu“ ryšį apsimoka, nes pagreitinamas žemėnaudos atributų įvedimas – operatoriui nebereikia kiekvieną kartą įvesti viso žemėnaudos klasės pavadinimo, o pakanka įvesti tik skaitmeninį kodą. Taip ne tik pagreitinamas duomenų įvedimas, bet ir sumažinama klaidų tikimybė (pvz., rašybos klaidų).
- **daug su daug** – daug vienos lentelės įrašų gali būti susiję su daug kitos lentelės įrašų. Pavyzdžiui, viename ūkyje galima auginti kelių tipų daržoves, ir šių tipų daržoves galima auginti daugiau negu viename ūkyje.

Atributų lenteles galima susieti su erdviniais sluoksniais jungimo (*join*) ir siejimo (*relate*) funkcijomis. Šios funkcijos yra:

Jungimas

Jungiant lenteles, vienos lentelės laukai prijungiami prie kitos lentelės pagal ryšį, sudarytą abiemis lentelėms bendro atributo (pvz., identifikatoriaus) pagrindu. Jungimas paprastai naudojamas tada, kai norima prie erdvinio duomenų sluoksnio pridėti papildomų atributų. Duomenis galima jungti sluoksnio arba lentelės vaizde. Nors laukų, pagal kuriuos jungiama, pavadinimai neprivalo būti tapatūs, duomenų tipas būtinai turi būti tas pats. Pavyzdžiui, tekstas jungiamas su tekstu, skaitinės reikšmės – su kitomis skaitinėmis reikšmėmis. Jungimo funkcija tinka tada, kai lentelės jungiamos „vienas su vienu“ ir „daug su vienu“ ryšiais. „Vienas su daug“ ryšiams jungimas netinka, nes pagrindinės lentelės įrašui priskiriama tik pirmoji aptikta tinkama reikšmė.

Jei redaguojate duomenis, atminti, kad keisti galima tik įvesties lentelės stulpelius. Prijungtų stulpelių redaguoti neįmanoma. Jei pridami nauji laukai, jie bus pridėti prie jungtinės lentelės, o jungiamos lentelės liks nepakitusios. Prijungtus stulpelius galima naudoti skaičiuojant jungtinės lentelės stulpelių reikšmes.

Siejimas

Kaip jau minėjome, jungimas naudojamas sudarant „vienas su vienu“ ir „daug su vienu“ ryšius tarp sluoksnių ir lentelių. Siejimo funkcija tinka „vienas su daug“ ir „daug su daug“ ryšiams sudaryti, nes siejimas yra dvikryptis. Susietos lentelės yra sujungtos, bet fiziškai išlieka atskiros – duomenys pateikiami dirbant su sluoksnio atributais. Siejami elementai (pvz., laukai ir lentelės) saugomi atskirai, o siejant programa prašo įvesti ryšio pavadinimą. Siejimo pranašumas tas, kad galima susieti daug failų, bet tada gali sulėtėti ir susietų duomenų apdorojimas.

1.2 Vektorinių duomenų užklausos ir atranka

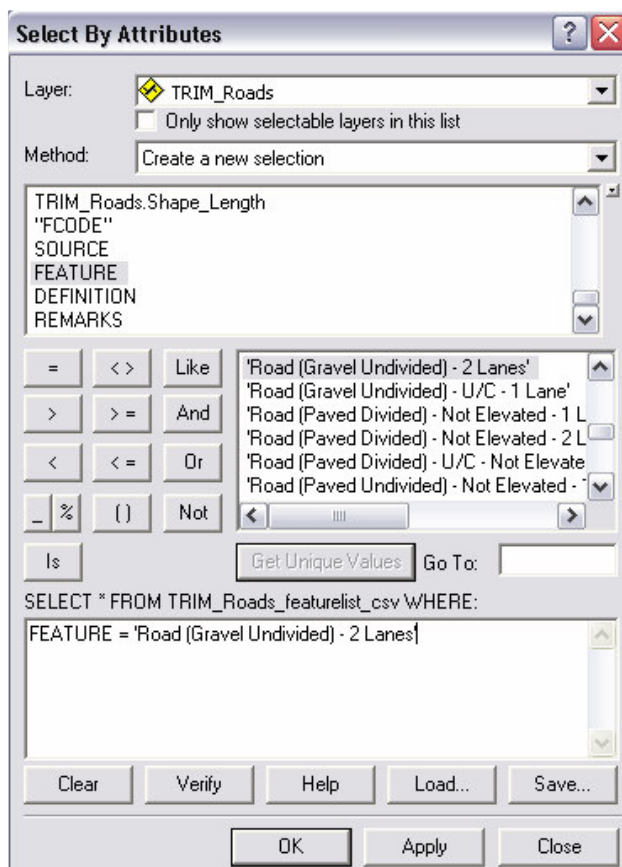
GIS duomenų užklausa yra pagrindinė funkcija ieškant reikiamų duomenų arba naujų erdvinio ryšių. Užklausa taip pat dažnai padeda sumažinti arba supaprastinti sudėtingus arba griozdiškus duomenų rinkinius. Jos palengvina sudėtingą interpretavimą ir analizę. Užklausa sukuria duomenų poaibius; paprastai yra du užklausų arba atrankos tipai – atributų (neerdvinės) ir erdvinės užklausos.

1.2.1 Atranka pagal atributą

Atributų užklausa – tai klausimai apie duomenų atributus (neerdvinės charakteristikos), pavyzdžiui, kiek kelių transporto infrastruktūros sluoksnyje yra dviejų juostų žvyrkeliai? Kadangi atributai yra tikra su elementais susieta informacija, geriausi atsakymai į GIS analizės klausimus dažnai būna atributų lentelėse. Struktūrizuotų užklausų kalba (SQL) yra standartinė sąsaja, kurioje atitinkantiems įrašams rasti (t. y. poaibiams atrinkti) naudojamos loginės išraiškos. SQL užklausų sintaksė įvairiuose programų paketuose skiriasi; toliau pateiktame pavyzdyje aprašyta, kaip tipinė užklausa atliekama *ArcMap* programa. Pavyzdyje aprašoma, kaip iš viso kelių duomenų rinkinio atrinkti tam tikro tipo kelius.

• Atrankos pagal atributus veiksmų tvarka

- *ArcMap* meniu juostoje spustelėkite *Selection* (pasirinkimas) ir pasirinkite *Select By Attributes* (išrinkti pagal atributus); arba atvertoje atributų lentelėje spustelėkite *Options* (pasirinktys) ir pasirinkite *Select By Attributes*.
- Atrankos pagal atributus dialogo lange spustelėkite sluoksnių (*Layer*) meniu ir pasirinkite sluoksnį su elementais, kuriuos norite atrinkti (1.7 pav.).
- Spustelėkite metodo meniu (*Method*) ir pasirinkite atrankos metodą (daugiau informacijos apie atrankos metodus rasite 1.2.3 skyriuje).
- Dukart spustelėkite (arba surinkite dialogo lange) norimą atributų lauko pavadinimą.
- Jei norite prie SQL išraiškos pridėti operatorių (pvz., „=“ arba „>“), spustelėkite atitinkamą operatoriaus mygtuką.
- Jei norite peržiūrėti pasirinkto lauko reikšmes, spustelėkite mygtuką *Get Unique Values* (gauti unikalios reikšmes).
- Kad pridėtumėte prie SQL išraiškos lauko reikšmę, dukart ją spustelėkite (arba surinkite klaviatūra konkrečią ieškomą reikšmę).
 - Atkreipkite dėmesį, kad skirtingo tipo duomenų užklausos sintaksė truputį skiriasi.
 - Užklausiama tekstinės reikšmės išskiriamos viengubomis kabutėmis (pvz., [FEATURE] = ,Tiltas‘)
 - Užklausiama skaitinių laukų reikšmės jokiais ženklais neišskiriamos (pvz., [PLOTAS] > 12.0)
- Paleiskite užklausą mygtuku *Apply* (vykdyti).



1.7 pav. Atrankos pagal atributą dialogo langas

- Atrinktų objektų skaičius bus parodytas kairiajame apatiniame *ArcMap* kampe. Žemėlapiu lange bus pažymėti visi su atrinktais atributais susiję elementai.

Loginiai (Bulio) operatoriai

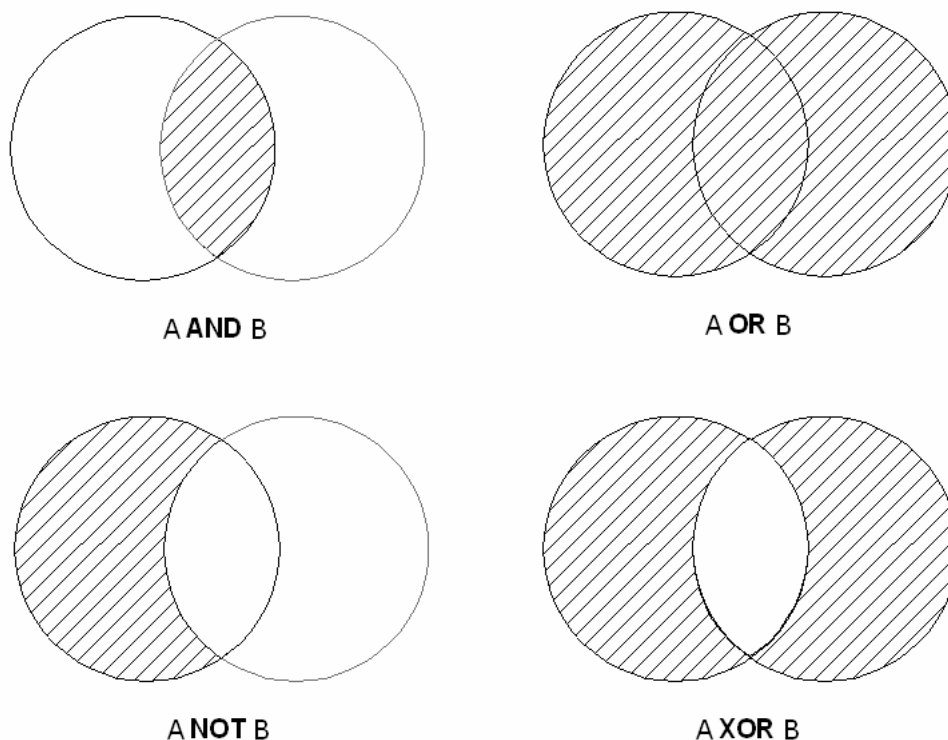
Bulio (arba loginiai) operatoriai naudojami sudaryti atrankos sąlygoms iš atskirų kriterijų. Sąlyga gali įgyti dvi reikšmes – būti teisinga arba klaidinga. Yra šie loginiai operatoriai:

- (=) lygu
- (>) daugiau
- (<) mažiau
- (>=) daugiau arba lygu
- (<=) mažiau arba lygu
- (<>) nelygu

Loginės funkcijos

Tarkime, kad norime atrinkti objektus pagal du arba daugiau kriterijų, pavyzdžiui, kiek dvijuosčių žvyrkelių buvo nutiesta 2002 metais? Tokiais atvejais atskiras užklausa galima sujungti taip, kad jos atsakytų į sudėtingesnę klausimą. Sudėtinių klausimų dalys jungiamos loginėmis funkcijomis – AND, OR, NOT, XOR (1.8 pav.):

- AND (IR) – turi būti tenkinami du arba daugiau sudėtinės užklauskos kriterijų;
- OR (ARBA) – turi būti tenkinamas bet kuris užklauskos kriterijus (arba abu);
- NOT (NE) – turi būti tenkinamas vienas užklauskos kriterijus, bet ne kitas;
- XOR (išskirtinis ARBA) – turi būti tenkinamas vienas arba kitas užklauskos kriterijus, bet ne abu kartu (t. y. tarpusavyje nesuderinami).

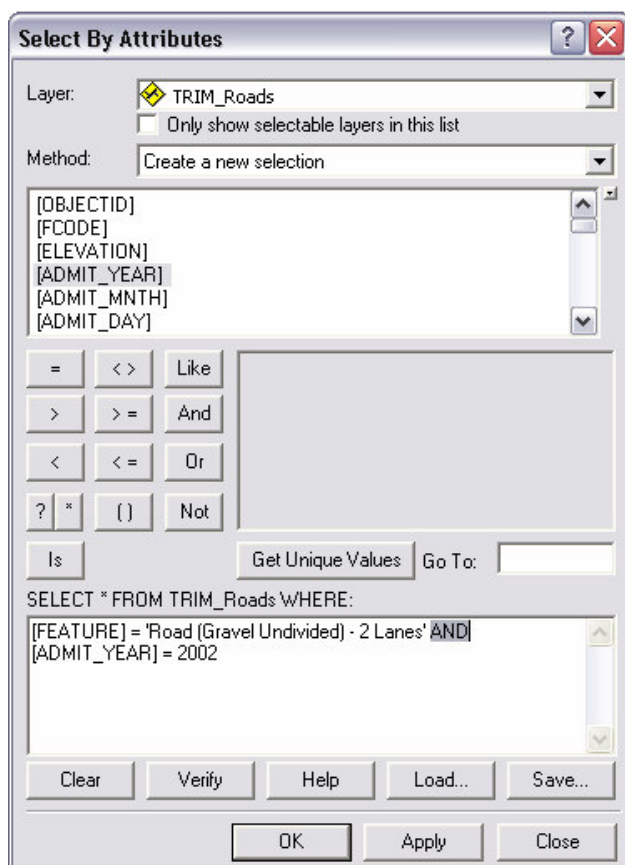


1.8 pav. Loginės funkcijos (Veno diagramos)

Kad klausimas būtų teisingas, reikia pasirinkti tinkamą loginę funkciją. Atsakymas į klausimą „Kiek kelių yra 2 juostų žvyreliai IR buvo nutiesti 2002 metais?“ skirsis nuo atsakymo į klausimą „Kiek kelių yra 2 juostų žvyreliai ARBA buvo nutiesti 2002 metais?“. Atsakyme į pastarąjį klausimą įrašų greičiausiai bus daugiau, nes įrašas turi atitikti tik vieną kriterijų, o ne abu.

• **Atrankos pagal atributus veiksmų tvarka (du arba daugiau kriterijų)**

- Atrankos pagal atributus dialogo lange dukart spustelėkite (arba surinkite klaviatūra) norimą atributų lauko pavadinimą, operatorių ir konkrečią ieškomą reikšmę.
- Pakartokite ankstesnįjį veiksma antrajam kriterijui. Atkreipkite dėmesį, kad užklauskos sujungtos IR (AND) loginė funkcija (1.9 pav.).
- Paleiskite užklausą mygtuku *Apply* (vykdyti).



1.9 pav. Atranka pagal atributą su loginėmis funkcijomis

1.2.2 Erdvinės užklausos ir ryšiai (atranka pagal vietą)

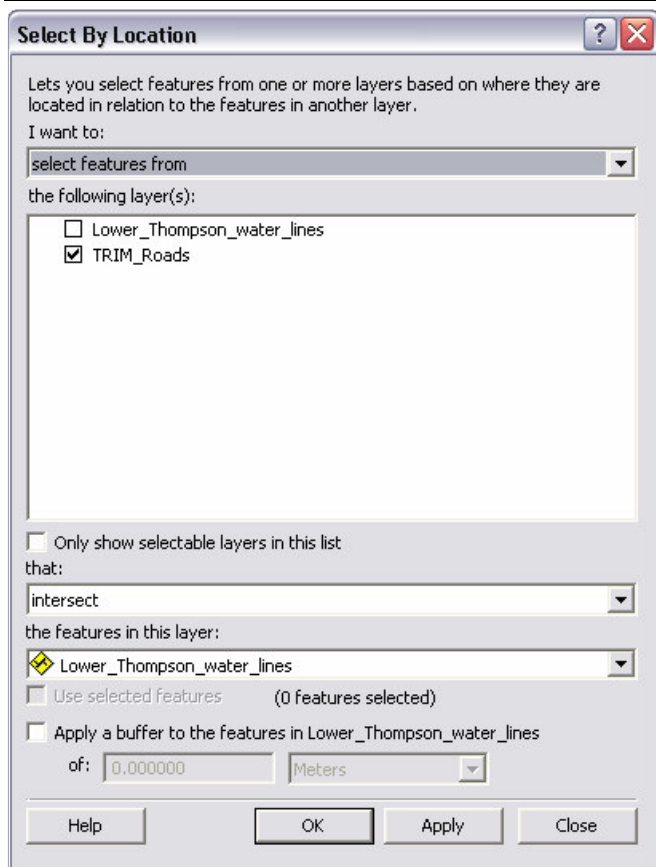
Erdvinėmis užklausomis klausiama, „kur“ erdvėje yra tam tikri objektai ir kaip jie susiję su kitais objektais. Šios užklausos leidžia atrinkti objektus pagal jų absoliučią padėtį ir (arba) vietą kitų sluoksnių objektų atžvilgiu. Pavyzdžiui, erdvine užklausa galima atrinkti visas tam tikro vandens baseino mėginių ėmimo vietas. Paprasčiausia erdvinė užklausa yra objekto pažymėjimas pele žemėlapiu vaizde.

• **Objektų atrankos žymekliu veiksmų tvarka.** Paprasčiausias būdas atrinkti ir pažymėti objektus – spustelėti pele vieną objektą; jei reikia atrinkti kelis vienu metu – *ArcMap* „elementų žymėjimo“ įrankiu aplink dominančią sritį nubrėžti stačiakampį. Veiksmai:

- Įsitikinkite, kad pasirinkimo skyrelyje (*Selection*) *ArcMap* turinio apačioje pažymėtas reikiamas sluoksnis.
- Spustelėkite įrankių juostos mygtuką *Select Features* (pasirinkti elementus).
- Žemėlapiu lange spustelėkite vieną objektą arba nubrėžkite stačiakampį aplink dominančių objektų grupę.
- Objektai bus tuo pat metu atrinkti ir pažymėti ir žemėlapiu lange, ir atributų lentelėje.

• **Atrankos pagal vietą veiksmų tvarka.** Daugiau galimybių negu paprasčiausias objektų pažymėjimas turi užklausa, nagrinėjanti erdvinius ryšius tarp skirtingų sluoksnių objektų (pvz., kur yra upės kertantys keliai?).

- Spustelėkite *ArcMap* meniu juostos mygtuką *Selection* (pasirinkimas) ir pasirinkite *Select By Location* (išrinkti pagal vietą).
- Atrankos pagal vietą lange spustelėkite išskleidžiamojo meniu rodyklę ir nurodykite atrankos metodą (daugiau informacijos apie atrankos metodus rasite 1.2.3 skyriuje) (1.10 pav.).
- Pažymėkite sluoksnį(-ius), kurių objektus norite atrinkti (paskirties sluoksnį).
- Išskleidžiamajame meniu nurodykite, kokio erdvinio ryšio ieškote:
 - *Intersect* (kertasi) – randa bet kokius objektus, turinčius bendrą dalių su pradiniais objektais;
 - *Are within a distance of* (yra ne toliau negu) – randa bet kokius objektus, esančius ne toliau negu nurodytu buferio atstumu nuo pradinių objektų;
 - *Completely contain* (visiškai apima) – randa bet kokius objektus, visiškai apimančius pradinio sluoksnio objektus.
 - *Are completely within* (visiškai patenka) – randa bet kokius objektus, visiškai patenkančius į pradinio sluoksnio objektą (-us);
 - *Have their centre in* (centras yra viduje) – randa bet kokius objektus, kurių centroidas patenka į pradinį objektą;
 - *Share a line segment with* (turi bendrą linijos atkarpą) – randa bet kokius objektus, kurie turi nors dvi gretimas viršūnes, sutampančias su pradinio objekto viršūnėmis;
 - *Touch the boundary of* (ribojasi su) – randa bet kokius objektus, kurie liečiasi su pradinio sluoksnio objektų ribomis;
 - *Are identical to* (identiški) – randa bet kokius objektus, tiksliai lygius (turinčius tas pačias viršūnes) pradinio sluoksnio objektui;
 - *Are crossed by the outline of* (kertami kito kontūro) – randa bet kokius objektus, turinčius bent vieną kraštinę, viršūnę arba galinį tašką, priklausančius ir pradinio sluoksnio objektui;
 - *Contain* (apima) – randa bet kokius objektus, apimančius pradinio sluoksnio objektą;
 - *Are contained by* (patenka į) – randa bet kokius objektus, patenkančius į pradinio sluoksnio objektą(-us).
- Apatiniame išskleidžiamajame meniu nurodykite, sluoksnį, kurį norite susieti su pradiniu sluoksniu.
- Paleiskite užklausą mygtuku *Apply* (vykdyti).
- Kaip ir atrenkant pagal atributus, atrinkti objektai bus pažymėti žemėlapiu lange. Be to, tuo pat metu atributų lentelėje bus atrinkti ir pažymėti atrinktų objektų atributai.



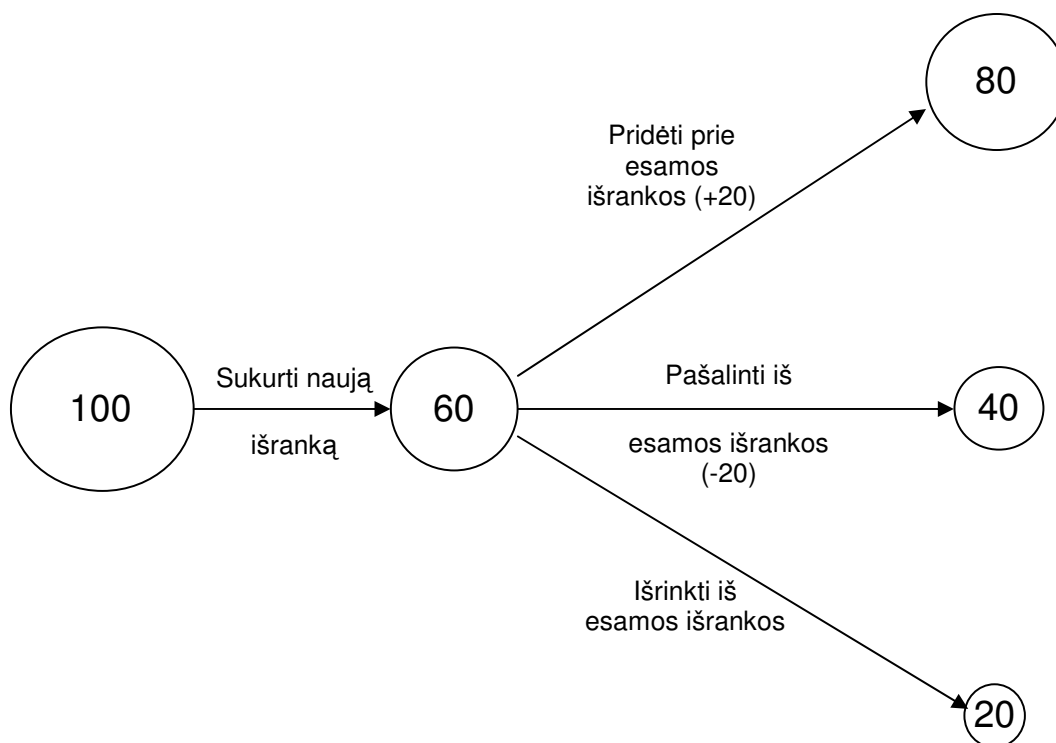
1.10 pav. Atrankos pagal vietą dialogo langas

1.2.3 Atrankos metodai

Galima pasirinkti kelis atributų ir erdviųjų išrankų kūrimo ir keitimo būdus (1.2 lentelė). Nors atrankos pagal atributą ir pagal vietą metodų pavadinimai skiriasi, gaunamų išrankų tipai iš principo tie patys.

1.2 lentelė. Atrankos pagal atributą ir pagal vietą metodai

Atranka pagal atributą	Atranka pagal vietą	Rezultatas
<i>create a new selection</i> (sukurti naują išranką)	<i>select features from</i> (išrinkti elementus iš)	sukuria naują išranką
<i>add to current selection</i> (pridėti prie esamos išrankos)	<i>add to the currently selected features in</i> (pridėti prie dabar išrinktų elementų)	prideda įrašus prie ankstesne užklausa gautos išrankos
<i>remove from current selection</i> (pašalinti iš esamos išrankos)	<i>remove from the currently selected features in</i> (pašalinti iš dabar išrinktų elementų)	pašalina įrašus iš ankstesne užklausa gautos išrankos
<i>select from current selection</i> (išrinkti iš esamos išrankos)	<i>select from the currently selected features in</i> (išrinkti iš dabar išrinktų elementų)	iš ankstesne užklausa gautos išrankos išrenka mažesnę įrašų poaibį



1.11 pav. Atrankos metodai

Sukūrus išranką, galima ją perjungti – sukeisti išrinktus ir neatrinktus poaibius. Atkreipkite dėmesį, kad reikia sekti, kokio tipo atranką šiuo metu vykdate, priešingu atveju galite gauti neteisingus rezultatus.

• Atrinkto ir neatrinkto poaibių perjungimo veiksmų tvarka

- Atributų lentelėje spustelėkite *Options* (parinktys) ir pasirinkite *Switch Selection* (perjungti išranką). Visi prieš tai atrinkti įrašai taps neatrinkti, o neatrinkti įrašai taps atrinkti.
- Jei norite išvalyti išranką, kad visi įrašai taptų neatrinkti, spustelėkite *Options* ir pasirinkite *Clear Selection* (išvalyti išranką).

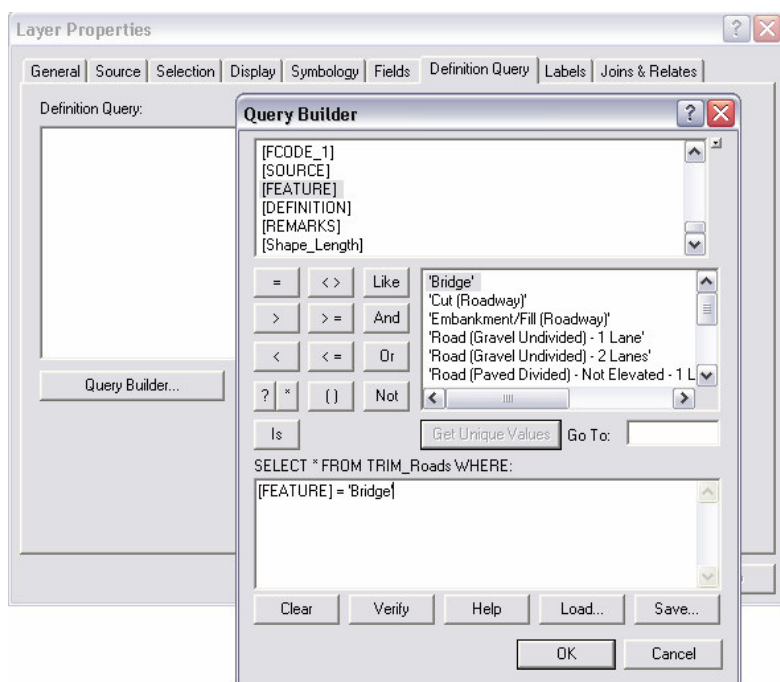
1.2.4 Ribojančios užklausos

Ribojančiomis užklausomis žemėlapiu lange (ir susijusiose atributų lentelėse) galima parodyti objektus, kurie turi tam tikras atributų reikšmes. Pavyzdžiui, atlikdamas įprastinę apžiūrą, inžinierius gali norėti ekrane matyti ir tikrinti ne visus transporto infrastruktūros sluoksnio elementus, o tik tiltus.

• Ribojančios užklausos įvedimo veiksmų tvarka

- Turinyje dešiniuoju pelės klavišu spustelėkite norimą sluoksnį ir pasirinkite *Properties* (savybės).
- Spustelėkite ribojančių užklausų (*Definition Query*) skyrelį.
- Klaviatūra surinkite SQL išraišką arba spustelėkite užklausų kūrimo (*Query Builder*) mygtuką (1.12 pav.).
 - Dukart spustelėkite (arba surinkite dialogo lange) norimą atributų lauko pavadinimą.

- Atkreipkite dėmesį, kad skirtingų įvesties formatų elementų klasės išskiriamos skirtingai.
- Užklaudami failinės geoduomenų bazės, SHAPE failo arba Arc/INFO sluoksnio elementų klasę, išskirkite laukų pavadinimus dvigubomis kabutėmis (pvz., "PLOTAS").
- Užklaudami asmeninės geoduomenų bazės elementų klasę, laukų pavadinimus išskirkite laužtiniais skliausteliais (pvz., [PLOTAS]).
- Užklaudami SDE elementų klasę, laukų visai neišskirkite (pvz., PLOTAS).
- Jei norite prie SQL išraiškos pridėti operatorių (pvz., „="), spustelėkite jo mygtuką.
- Jei norite peržiūrėti pasirinkto lauko reikšmės, spustelėkite mygtuką *Get Unique Values* (gauti unikalias reikšmes).
- Kad pridėtumėte prie SQL išraiškos lauko reikšmę, dukart ją spustelėkite (arba surinkite klaviatūra konkrečią ieškomą reikšmę).
 - Atkreipkite dėmesį, kad skirtingo tipo duomenų užklaudos sintaksė truputį skiriasi.
 - Užklaudamos tekstinės reikšmės išskiriamos viengubomis kabutėmis (pvz., [FEATURE] = 'Tiltas')
 - Užklaudamos skaitinių laukų reikšmės jokiais ženklais neišskiriamos (pvz., [PLOTAS] > 12.0)
- Spustelėkite *OK*.

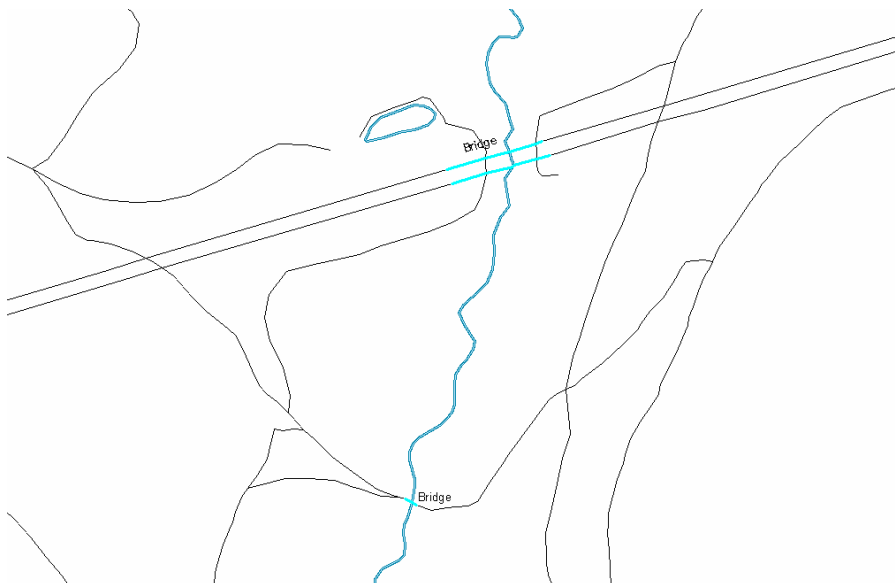


1.12 pav. Ribojančios užklaudos dialogo langas

- Paleiskite užklausą mygtuku *Apply* (vykdyti).
- Jei norite vėl matyti visus objektus, pašalinkite ribojančią užklausą.

1.2.5 Išrankų peržiūra

Kai užklausa atlikta ir išranka sukurta, gautą naują duomenų poaibį verta peržiūrėti. Pagal atributą arba vietą atrinkti objektai ArcMap žemėlapiu lange yra pažymėti žydra spalva (1.13 pav.).



Bridge	Tiltas
--------	--------

1.13 pav. Išrankos peržiūra žemėlapiu lange

Susijusioje atributų lentelėje galima peržiūrėti ir išanalizuoti atrinktus neerdvinius duomenis. Šie įrašai atributų lentelėje bus pažymėti, bet galima padaryti, kad būtų rodomi tik atrinkti įrašai.

• Išrankos peržiūros atributų lentelėje veiksmų tvarka

- Turinyje dešiniuoju pelės klavišu spustelėkite norimą sluoksnį ir pasirinkite *Open Attribute Table* (atverti atributų lentelę).
- Spustelėkite atributų lentelės apačioje esantį mygtuką *Selected* (atrinktus), ir ekrane bus rodomi tik atrinkti objektai (įrašai) (1.14 pav.).

Selected Attributes of TRIM_Roads

StudyArea	F_CODE_1	SOURCE	FEATURE	DEFINITION	REMARKS	Shape_Length
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	7.080038
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	10.607994
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	12.509232
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	12.175002
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	12.022175
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	9.199101
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	9.187373
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	6.999127
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	10.1742
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	4.992342
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	5.85888
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	4.48086
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	4.534489
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	14.874354
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	8.896968
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	6.108204
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	4.339004
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	14.716487
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	3.002501
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	16.524737
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	25.635126
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	6.291106
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	30.13114
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	29.927496
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	38.032356
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	25.635071
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	43.165524
Lower Thompson	DD93250000	TRIM	Bridge	A structure erected along a travelled route to span a depression or obstacle.	<Null>	14.461672

Record: 1 Show: All Selected Records (28 out of *2000 Selected) Options

1.14 pav. Išrankos peržiūra atributų lentelėje

1.2.6 Atrinktų duomenų kopijavimas (išskyrimas)

Jeigu tam tikrą duomenų rinkinio dalį norite išanalizuoti išsamiau, gali būti naudinga atskirti šį duomenų poaibį nuo pradinio duomenų rinkinio. Tai daroma išskiriant (eksportuojant) atrinktus objektus į naują sluoksnį arba atrinktus atributus į naują lentelę. Taip pat dažnai naudinga nukopijuoti atrinktus duomenis į kitą esamą sluoksnį arba lentelę, kad būtų galima sujungti šį pradinio failo poaibį su kitu duomenų rinkiniu.

• Atrinktų atributų eksportavimo į naują lentelę veiksmų tvarka

- Atverkite elementų klasės, kurios užklausa atliksite, atributų lentelę.
- Rankiniu būdu arba atrankos pagal atributus (*Select by Attributes*) funkcija sukurkite lentelės įrašų išranką.
- Spustelėkite *Options* (pasirinkty) ir pasirinkite *Export* (eksportuoti).
- Spustelėkite eksportavimo išskleidžiamąjį meniu (*Export*) ir pasirinkite *Selected Records* (atrinktus įrašus).
- Spustelėkite mygtuką *Browse* (naršyti) ir nurodykite esamą aplanką arba geoduomenų bazę, į kurią eksportuosite.
- Išsaugojimo tipo (*Save as type*) išskleidžiamajame meniu nurodykite lentelės, į kurią eksportuosite, tipą (pvz., geoduomenų bazės lentelė, dbase, tekstinis failas).
- Surinkite būsimos lentelės pavadinimą ir spustelėkite *Save* (išsaugoti), tada spustelėkite *OK*.
- Tada naują lentelę peržiūrėti ir jos duomenis analizuoti galima kitomis programomis. Pavyzdžiui, *MS Excel* arba *Access*. Tai naudinga ir sumuojant bei interpretuojant lentelių pavidalo duomenis.

• Atrinktų objektų eksportavimo į naują sluoksnį veiksmų tvarka

- Sukurkite objektų išranką.
- Turinyje dešiniuoju pelės klavišu spustelėkite sluoksnį su atrinktais objektais ir pasirinkite *Data → Export Data* (duomenys → eksportuoti duomenis).
- Norimą koordinatų sistemą nurodykite akutėmis (*radio buttons*).

- Spustelėkite mygtuką *Browse* (naršyti) ir nurodykite norimą geoduomenų bazę arba aplanką.
 - Išsaugojimo tipo (*Save as type*) išskleidžiamajame meniu nurodykite elementų klasės, į kurią eksportuosite, tipą (pvz., geoduomenų bazės elementų klasė arba SHAPE failas).
 - Surinkite būsimą slauksnio pavadinimą ir spustelėkite *Save* (išsaugoti), tada spustelėkite *OK*.
 - Tada naują elementų klasę galima įkelti į *ArcMap* tolesnei analizei.
- **Objektų kopijavimo iš vieno slauksnio į kitą esamą slauksnį veiksmų tvarka**
- *ArcMap* programoje pradėkite redaguoti slauksnį – spustelėkite *Editor* (redaktorius) ir pasirinkite *Start Editing* (pradėti redaguoti), prieš tai įsitikinę, kad reikiamas slauksnis yra paskirties slauksnis.
 - Sukurkite pradinio slauksnio objektų išranką.
 - Spustelėkite įrankių juostos mygtuką *Copy* (kopijuoti).
 - Spustelėkite įrankių juostos mygtuką *Paste* (įdėti).
 - Atrinkti pradinio slauksnio objektai bus nukopijuoti į paskirties slauksnį.
 - Spustelėkite *Editor* ir pasirinkite *Save Edits* (išsaugoti pakeitimus), tada spustelėkite *Stop Editing* (baigti redaguoti).

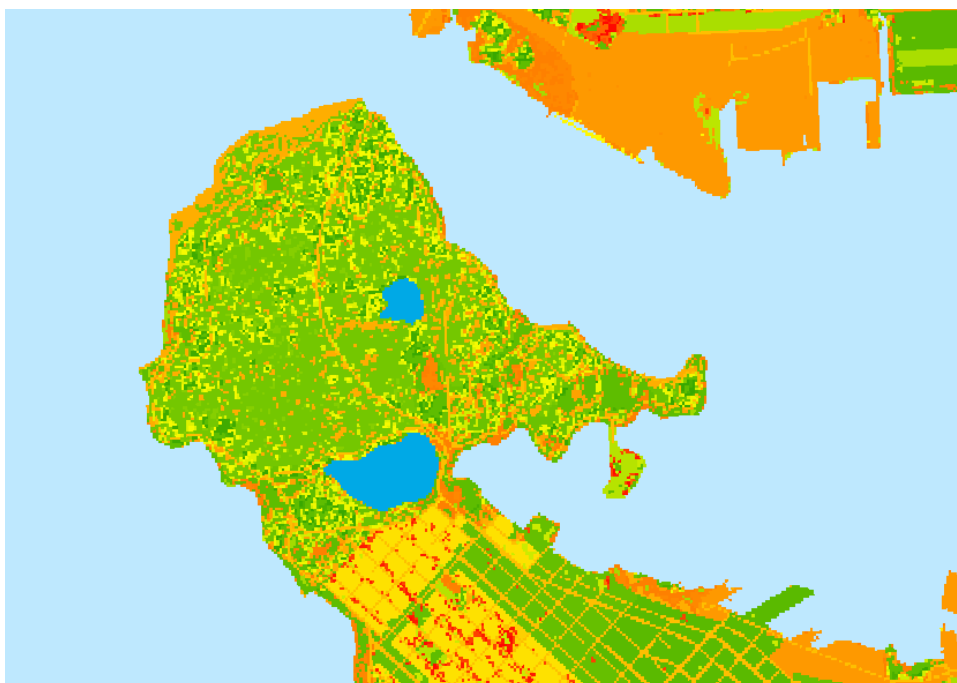
1.3 Rastrinių duomenų užklausos ir atranka

Kaip ir vektoriniams sluoksniams, naudodami rastrinius duomenis taip pat galima formuoti užklausas pagal atributą arba vietą erdvėje ir gauti konkretesnei analizei reikalingos informacijos poaibius.

1.3.1 Rastrinių duomenų išskyrimas

Atranka pagal atributą

Kai kuriuose rastriniuose duomenyse taip pat yra atributų lentelių. Tai gali būti, pavyzdžiui, palydovinis vaizdas, klasifikuotas rastriniam žemėnaudų žemėlapiui sukurti. Jei yra rastro atributų lentelė, galima atverti atrankos pagal atributą (*Select by Attribute*) dialogą ir panašiu į skirtą vektoriniams duomenims būdu atrinkti rastro gardeles. Rastrų atveju ši funkcija gali tikti nustatant ir sumuojant ežerų užimamą plotą žemėnaudos duomenų rinkinyje (1.15 pav.).



1.15 pav. Atranka pagal rastro atributą

$$\begin{aligned} \text{Plotas} &= \text{gardelių skaičius su žemės danga} = \text{'Ežeras'} \times \text{gardelės plotas} \\ &= 53\,619 \times 225 \text{ m}^2 = 12\,064\,275 \text{ m}^2 \text{ (1 206 ha)} \end{aligned}$$

• Atrankos pagal atributus veiksmų tvarka

- Rastro atributų lentelėje spustelėkite *Options* (parinktys) ir pasirinkite *Select By Attributes* (atrinkti pagal atributus).
- Atrankos pagal atributus lange spustelėkite metodo (*Method*) išskleidžiamąjį meniu ir pasirinkite atrankos metodą (daugiau informacijos apie atrankos metodus rasite 1.2.3 skyriuje).
- Dukart spustelėkite (arba surinkite dialogo lange) norimą atributų lauko pavadinimą (rastrų atveju tai paprastai bus koduotos „reikšmės“ (*'value'*) laukas, atitinkantis tikrąją savybę, pavyzdžiui, altitudę arba žemėnaudą).

- Jei norite prie SQL išraiškos pridėti operatorių (pvz., „=“ arba „>“), spustelėkite atitinkamą operatoriaus mygtuką.
- Jei norite peržiūrėti pasirinkto lauko reikšmes, spustelėkite mygtuką *Get Unique Values* (gauti unikalias reikšmes).
- Kad pridėtumėte prie SQL išraiškos lauko reikšmę, dukart ją spustelėkite (arba surinkite klaviatūra konkrečią ieškomą reikšmę – išskirkite ją viengubomis kabutėmis).
- Paleiskite užklausą mygtuku *Apply* (vykdyti).

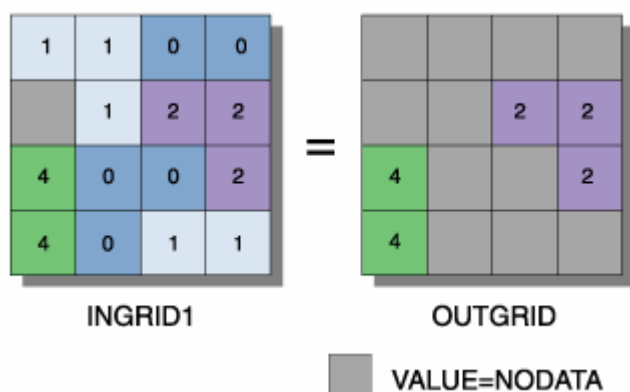
Išskyrimas pagal atributą

Rastro gardelės išskiriamos užklausa su „kur“ (*where*) sąlyga, kurią įvykdžius gaunamas naujas rastras. „Kur“ sąlyga – tai sąlyginis sakiny, nustatantis norimus užklausos kriterijus. Įsivaizduokite biologą, modeliuojantį kalnų ožkų buveines. Šie gyvūnai gyvena tik stačiuose šlaituose. Paprasčiausias būdas nustatyti galimas jų buveinių vietas – paimti nuolydžio rastrą ir išskirti tik tas gardeles, kurių nuolydis didesnis arba lygus 20 proc. Šiuo atveju sąlyga būtų `[NUOLYDIS] >= 20` (laužtiniai skliausteliai [] apie nuolydžio rastrą – įprastinė *ArcMap* sintaksė dirbant su rastrais).

• Išskyrimo pagal atributus veiksmų tvarka (*Extract by Attributes* įrankis)

- Įrankių juostoje spustelėkite raudoną įrankinės mygtuką ir atverkite ArcToolbox įrankinę.
- Pasirinkite išskyrimo pagal atributus įrankį (*Spatial Analyst Tools* → *Extraction* → *Extract by Attributes*).
- Norimoms gardelės reikšmėms atskirti sukurkite sakinį su „kur“ (*where*) sąlyga.
- Gautame rastre užklausos kriterijus tenkinančių gardelių reikšmės liks tos pačios kaip ir pradiniam duomenų rinkinyje, o kriterijų netenkinančioms gardelėms bus priskirta reikšmė NoData (1.16 pav.).

Select (NuolydžioRastras, ,NUOLYDIS >= 2')*



- 2-a nuolydžio klasė = 20 proc., 4-a = 40 proc.

INGRID1	ĮVESTIES RASTRAS 1
OUTGRID	IŠVESTIES RASTRAS
VALUE=NODATA	REIKŠMĖ=NODATA

1.16 pav. Išskyrimas pagal rastro atributą

Kaip ir vektorinėse užklausose, rastrinėse užklausose du arba daugiau kriterijų (arba rastrų) į vieną loginę išraišką galima sujungti loginėmis funkcijomis (AND, OR, NOT). Paprastai šios sudėtinės išraiškos naudojamos keliems rastriniams duomenų rinkiniams sujungti į vieną rastrą. Pavyzdžiui, anksčiau minėtą kalnų ožkų buveinės modelį galima pakeisti įtraukiant aukščių rastrą – buveinėms tinka tik vietos, kurių nuolydis didesnis arba lygus 20 proc. IR aukštis didesnis negu 1000 metrų. Šiuo atveju sakinyss būtų „([NUOLYDIS] >= 20) AND ([ALTIT] > 1000)“.

Atranka pagal vietą

Rastrų atveju atrenkant pagal vietą galima išskirti gardeles pagal vietą ir ištirti erdvinius ryšius tarp nesugretinamų sluoksnių. Tarkime, kad miesto projektuotojas nori susumuoti miško dangos plotą miesto teritorijoje – bet ne visos, o tik tam tikrų miesto parkų. GIS gali palengvinti šią užduotį – mišku apželdintiems miesto plotams iš viso miesto žemėnaudos rastro išskirti galima panaudoti parkų poligonų sluoksnį. Duomenų apribojimas iki parkų teritorijos padės susumuoti duomenis taip, kaip to reikia miesto projektuotojui.

Yra įvairių būdų rastro gardelėms pagal vietą išskirti – daugumoje jų rastrinio duomenų rinkinio gardelėms ar jų grupėms atrinkti ar atmesti naudojama tam tikra geometrinė figūra.

Išskyrimas trafaretu (*Extract by Mask*) – leidžia išskirti rastrinius duomenis pagal poligonų elementų klasę. Poligonas naudojamas kaip trafaretas – atrenkamos tik į jo vidų patenkančios gardelės.

Išskyrimas figūromis

Išskyrimas taškais (*Extract by Points*) – atrenkamos tik pateikto koordinačių (x ir y reikšmių) sąrašo taškuose esančios gardeles.

Išskyrimas apskritimu (*Extract by Circle*) – atrenkamos gardelės, patenkančios (arba nepatenkančios) į nurodyto spindulio apskritimą su centru nurodytame taške.

Išskyrimas poligonu / stačiakampiu (*Extract by Polygon/Rectangle*) – gardelių atrankos sritis nurodoma koordinačių sąrašu, ir išskiriamos į šią sritį patenkančios arba nepatenkančios rastro gardelės.

1.3.2 Perklasifikavimas

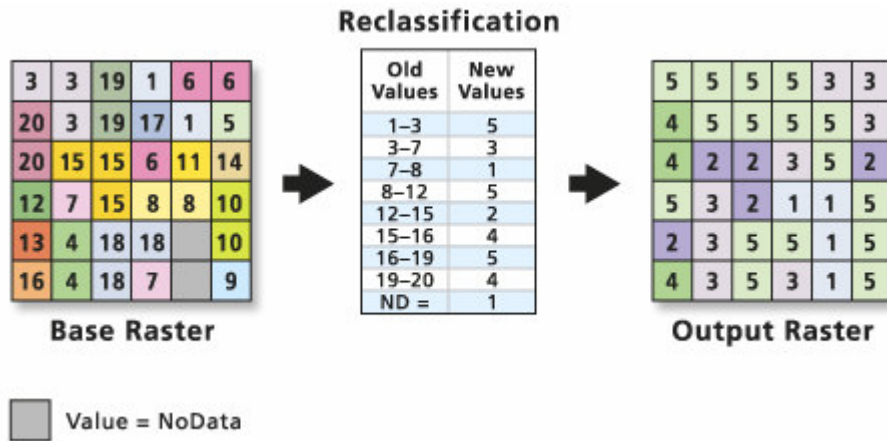
Rastrinio duomenų rinkinio duomenis supaprastinti arba agreguoti galima perklasifikuojant; perklasifikavimas dar vadinamas perkodavimu arba transformavimu. Pavyzdžiui, jei rastriniame duomenų rinkinyje yra 10 medžių rūšių, o jūs norite apjungti visas rūšis į vieną klasę, tam tinka perklasifikavimo funkcija.

• Rastrinio duomenų rinkinio perklasifikavimo veiksmų tvarka

- Įjunkite *Spatial Analyst* modulį: meniu juostoje spustelėkite *Tools* (įrankiai) ir pasirinkite *Extensions* (moduliai). Pažymėkite *Spatial Analyst* modulį varnele ir spustelėkite *Close* (užverti).
- Pridėkite *Spatial Analyst* įrankių juostą: dešiniuoju pelės klavišu spustelėkite bet kurioje įrankių juostos vietoje ir kontekstiniame meniu pasirinkite *Spatial Analyst*.
- Sluoksnių (*Layer*) išskleidžiamajame meniu pasirinkite paskirties rastrą (pvz., Žemėnauda).
- Spustelėkite *Spatial Analyst* ir pasirinkite *Reclassify...* (perklasifikuoti...).
- Dešiniajame perklasifikavimo dialogo lango stulpelyje įveskite naujas išvesties rastro reikšmes.

- Paleiskite perklasifikavimą mygtuku *OK*.

Rastro supaprastinimo grupuojant vienodų savybių gardoles pavyzdys (pvz., apjungiant tolydžias nuolydžio reikšmes į kelias nuolydžio klases [0–10 proc. = 1, 11–20 proc. = 2, ir t. t.]) pateiktas 1.17 pav.



Reclassification	Perklasifikavimas
Old values	Senos reikšmės
New values	Naujos reikšmės
Base Raster	Pradinis rastras
Output Raster	Išvesties rastras
Value = NoData	Reikšmė = NoData

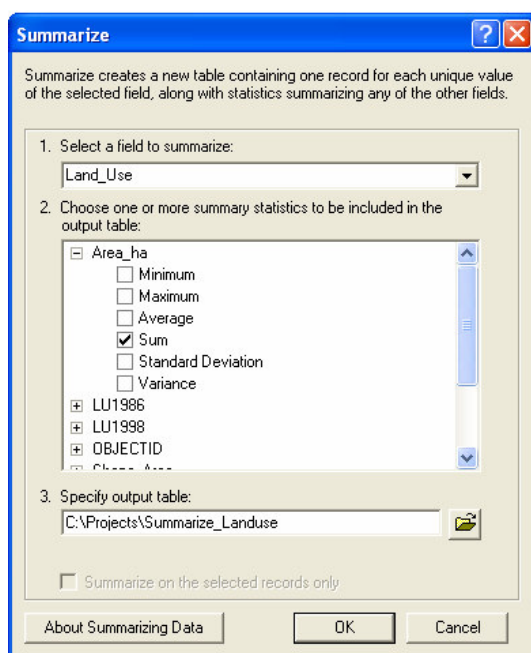
1.17 pav. Perklasifikavimas apjungiant reikšmes į grupes

1.4 Duomenų sumavimas ir interpretavimas

1.4.1 Duomenų sumavimas

Sumavimo funkcija leidžia klasifikuoti duomenis pagal atributą. Kitaip tariant, duomenis galima organizuoti pagal konkrečius reikalavimus. Sumuojant galima apskaičiuoti suvestinę duomenų statistiką (kiekį, vidurkį, mažiausią ir didžiausią reikšmes). Pavyzdžiui, žemėnaudos duomenų rinkinį iš šimtų poligonų, kuriame gali būti dešimt žemėnaudos klasių, galima susumuoti ir gauti visą kiekvienos klasės užimamą plotą. Šiuo atveju būtų sukurta nauja lentelė, kurios viename stulpelyje būtų išvardyti žemėnaudos tipai, o kitame pateiktos visų kiekvienos klasės poligonų plotų sumos.

Jei norite sumuoti duomenis *ArcMap* programa, atverkite atributų lentelę ir dešiniu juo pelės klavišu spustelėkite pavadinimą lauko, pagal kurį ketinate grupuoti sumuodami. Pasirinkite (spustelėkite kairiuoju pelės klavišu) sumavimo funkciją. Atsivers sumavimo dialogo langas (1.18 pav.). Pasirinkto lauko (*Select field to summarize*) langelyje bus jūsų pasirinkto lauko pavadinimas. Jei norite pasirinkti kitą lauką, tiesiog išskleiskite sąrašą ir pasirinkite vieną iš pateiktų laukų pavadinimų. Antruoju veiksmu dialogo lange reikia nurodyti suvestinę statistiką, kuri bus pateikta išvesties lentelėje. Šiame žemėnaudos pavyzdyje mums reikia susumuoti visą kiekvienos žemėnaudos klasės užimamą plotą, taigi pažymėkite (kairiuoju pelės klavišu) sumos (*Sum*) langelį po *Area_ha* atributų lauko, kuriame saugomas kiekvieno poligono plotas hektarais, pavadinimu. Paskutinis veiksmas – nurodyti išvesties lentelės pavadinimą. Jei buvo atrinktų įrašų (pvz., užklauso rezultatas), sumavimo funkcija leidžia pasirinkti, ar sumuoti visus įrašus, ar tik atrinktuosius. Kai nustatysite visas parinktis, spustelėkite mygtuką *OK*; kai programa paklaus, ar pridėti naująją lentelę, spustelėkite *Yes* (taip).

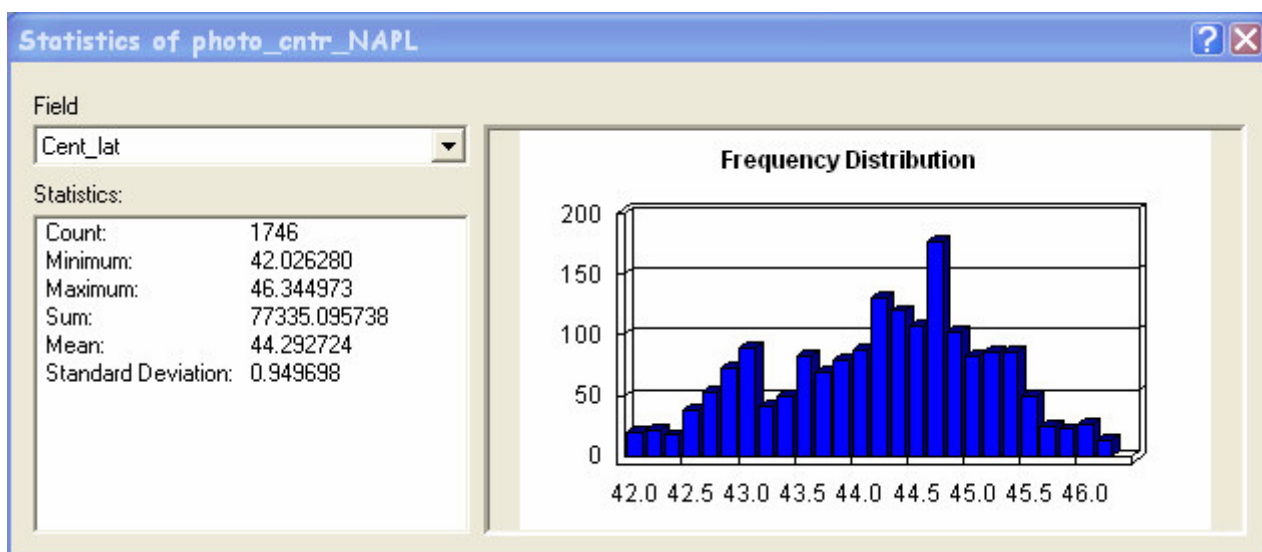


1.18 pav. Sumavimo dialogo langas

1.4.2 Statistika

GIS galima apskaičiuoti skaitinių laukų turinį apibūdinančią statistiką. Galima sužinoti įrašų skaičių, mažiausią ir didžiausią lauko reikšmę, reikšmių sumą, vidurkį ir standartinį nuokrypį. Be to, dauguma sąsajų parodys histogramą – lauko reikšmių dažnių pasiskirstymą. Šio dialogo lange galima gauti ir kito lauko statistiką – išskleidžiamajame sąrašė reikia pasirinkti kito lauko pavadinimą.

Jei norite apskaičiuoti lauko statistiką *ArcMap* programa, atverkite elementų klasės atributų lentelę ir dešiniuoju pelės klavišu spustelėkite pavadinimą lauko, kurį ketinate apibūdinti. Sąrašė pasirinkus *Statistics* (statistika), atsivers dialogo langas (1.19 pav.). Atkreipkite dėmesį, kad statistika skaičiuojama tik skaitiniams laukams.

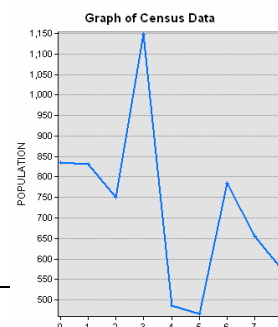


1.19 pav. Statistikos rezultatų pavyzdys

1.4.3 Diagramos

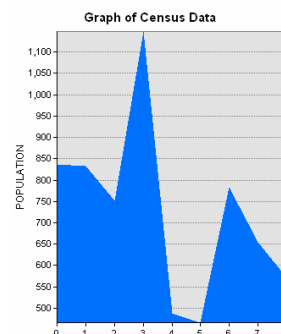
Diagramos apibendrina duomenis vaizdiniu formatu, kuris dažnai suprantamesnis negu lentelių pavidalo duomenys, kadangi duomenis vizualizuoti galima įvairiais būdais. Žemėlapiu maketuose diagramomis galima vaizduoti papildomą su žemėlapiu susijusią informaciją arba pateikti tą pačią informaciją kitu būdu. Anksčiau aptartame pavyzdyje apie žemėnaudą galėtume sukurti, pavyzdžiui, žemėlapi, kuriame kiekviena žemėnaudos klasė būtų pavaizduota kita spalva ir pateikti viso kiekvienos klasės ploto diagramą. Žemėlapiu langas ir diagrama papildys vienas kitą, ir skaitytojas gaus daugiau informacijos apie vaizduojamus objektus. Yra daug įvairių diagramų tipų – ir dvimačių, ir trimatinių. Kiekvienas diagramos tipas turi rodymo parametrus, kuriuos galima suderinti atsižvelgiant į vaizduojamus duomenis ir norimą jų pateikimo būdą. Toliau aprašyti įvairūs diagramų tipai ir pateikti keli įprastinių diagramų pavyzdžiai.

- **Linijinės (line) diagramos** – vaizduoja duomenis x, y koordinatų sistemoje. Diagramoje gali būti viena arba daugiau linijų. Linijinės diagramos tinka vaizduoti duomenų tendencijas tolydžioje skalėje. Šio tipo diagramomis galima aiškiai pavaizduoti gyventojų tankio arba bendrojo vidaus produkto (BVP) kitimą per kelis metus.



- **Polinės (*polar*) diagramos** – panašios į linijines, bet vaizduoja kampinius duomenis (išreikštus laipsniais arba radianais) polinėje koordinačių sistemoje. Jos tinka matematinių formulų rezultatams vaizduoti.

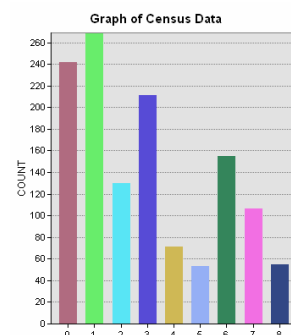
- **Plotinės (*area*) diagramos** – panašios į linijines tuo, kad jose taip pat vaizduojama viena arba daugiau linijų x, y koordinačių sistemoje, ir jos taip pat tinka duomenų tendencijoms vaizduoti. Plotinių diagramų spalvinimas leidžia geriau išryškinti kiekių skirtumus.



- **Skaidos (*scatter*) diagramos** – duomenų rinkinio atributų reikšmės vaizduojamos simboliais x, y (galbūt ir z) koordinačių sistemoje. Skaidos diagramomis galima aiškiai pavaizduoti kelis kintamuosius ir geriau tinka nagrinėjant kintamųjų tarpusavio ryšius. Pavyzdžiui, jei norime nustatyti, ar susijusios metinės pajamos su vidutine gyvenimo trukme, šiuos kintamuosius galime pavaizduoti skaidos diagrama.

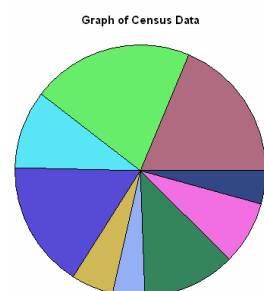
- **Burbulinės (*bubble*) diagramos** – panašios į skaidos diagramas, bet galima vaizduoti tris kintamuosius dvimatėje erdvėje. Šiuo atveju naudojami skirtingo dydžio simboliai (burbuliukai), kurių dydis proporcingas trečio kintamojo reikšmėms. Jei pajamas ir vidutinę gyvenimo trukmę lygintume ne skaidos, o burbuline diagrama, galėtume lyginti ir trečią svorio kintamąjį ir ieškoti galimos kintamųjų koreliacijos.

- **Juostinės ir stulpelinės (*bar and column*) diagramos** – dar vadinamos histogramomis. Jose duomenys grupuojami vienodais intervalais (į klases) ir pateikiami juostomis arba stulpeliais, vaizduojančiais kiekvienos klasės reikšmių skaičių (dažnį, tankį). Šio tipo diagramos tinka vaizduoti duomenų tendencijas (pvz., vidutinę mėnesio temperatūrą arba kritulių kiekį).



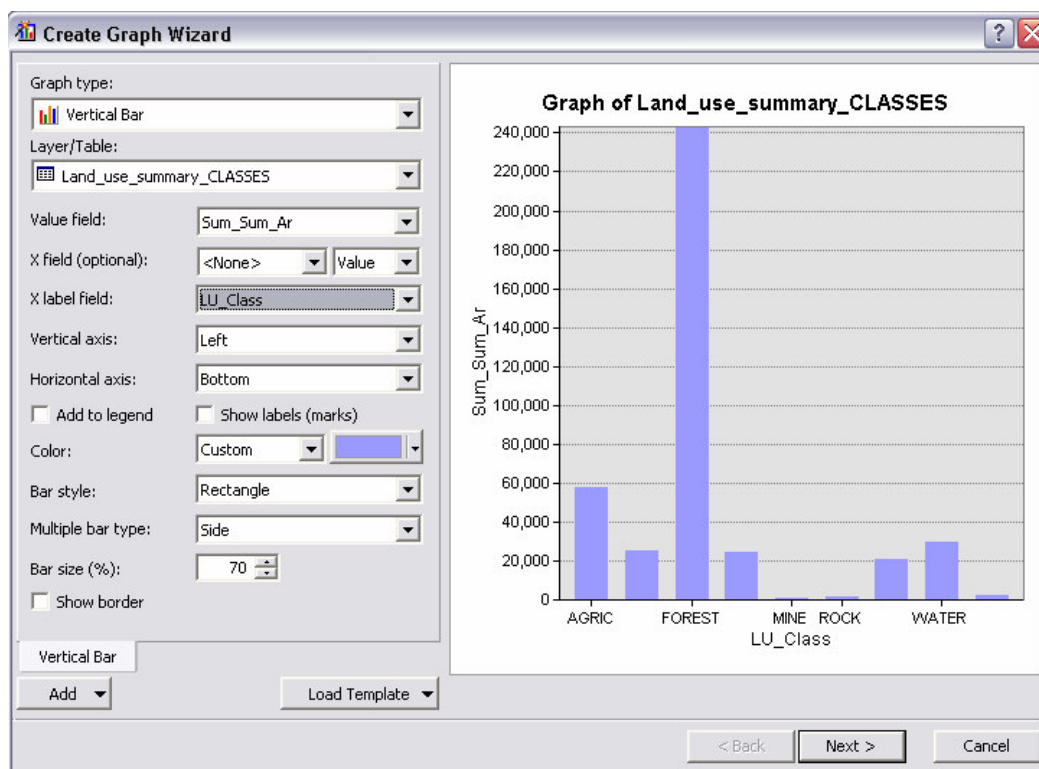
- **Didžiausios, mažiausios ir galutinės reikšmės (*high-low-close*) diagramos** vertikalia juosta vaizduoja kiekvienos x reikšmės y reikšmių diapazoną. Didžiausios ir mažiausios duomenų reikšmės vaizduojamos horizontaliomis linijomis. Šio tipo diagramomis galima vaizduoti kainų svyravimus per dieną – nurodyti pradinę, didžiausią, mažiausią ir galutinę kainą.

- **Skritulinės (*pie*) diagramos** – būna dvimatės ir trimatės. Duomenys vaizduojami skritulio segmentais, kurių dydis proporcingas atitinkamos reikšmės daliai visame rinkinyje. Šia diagrama galima aiškiai pavaizduoti įvairių žemėnaudos klasių ploto santykį.



Kuriant diagramą, reikia nurodyti vaizduojamus kintamuosius ir pasirinkti tokį diagramos tipą, kuriame vaizduojami duomenys būtų pateikti aiškiai. Jei diagramą norite sukurti ir

patalpinti žemėlapyje *ArcMap* programa, meniu juostoje pasirinkite *Tools (įrankiai)* ir diagramų parinktį (*Graph*) ir spustelėkite *Create* (sukurti). Pasirodys pirmasis diagramų vedlio (*Graph Wizard*) langas, kuriame programa paprašys nurodyti diagramos tipą (1.20 pav.). Kaip sukurti žemės dangos klasių stulpelinę diagramą, paaiškinsime pavyzdžiu.

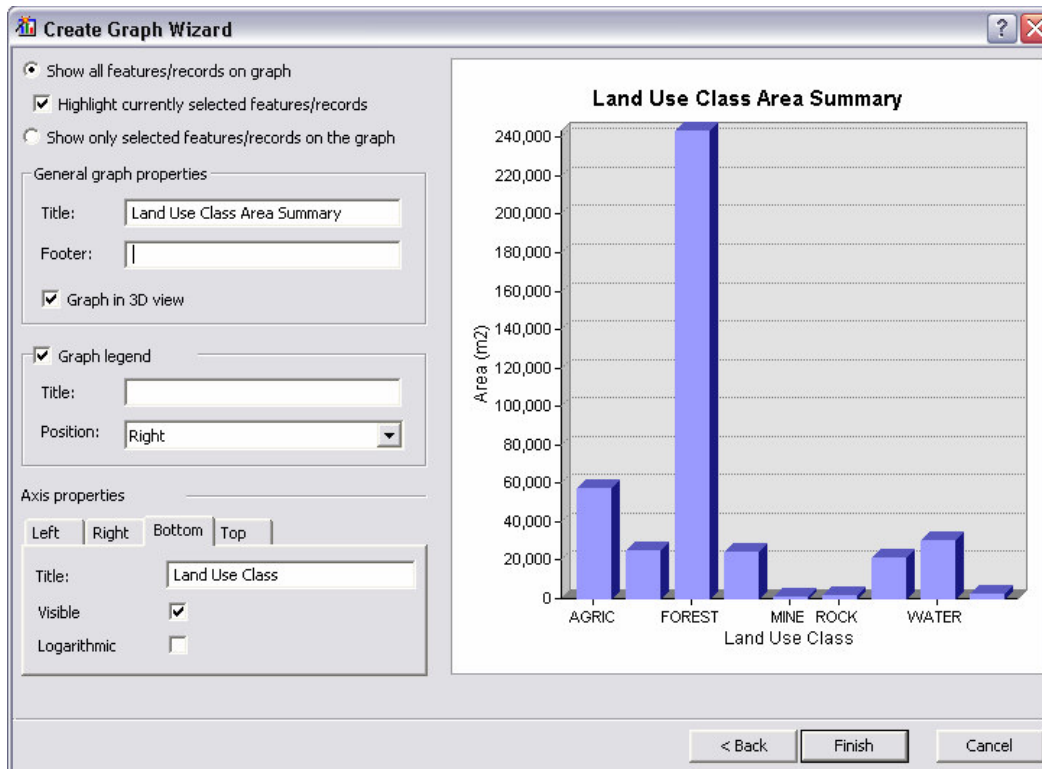


1.20 pav. 1 diagramų vedlio veiksmas (iš 2)

Diagramos vedliu pasirinkite reikiamą sluoksnį ir kurį nors stulpelinės diagramos stilių. Pasirinkę diagramos tipą ir (arba) stilių bei duomenų laukus, spustelėkite mygtuką *Next* (toliau). Atsivers antrasis vedlio langas (1.21 pav.). Galima nurodyti bendrąsias diagramos savybes, pavyzdžiui, diagramos pavadinimą bei x ir y ašių pavadinimus. Žymimaisiais langeliais galima nurodyti naudoti visus arba tik atrinktus failo įrašus.

- **Diagramos kūrimo veiksmų tvarka:**

1. Diagramos tipo (*Graph type*) išskleidžiamajame meniu pasirinkite kuriamos diagramos tipą.
2. Sluoksniu / lentelės (*Layer/Table*) išskleidžiamajame meniu pasirinkite diagrama vaizduojamą sluoksnį arba lentelę.
3. Reikšmių lauko (*Value field*) išskleidžiamajame meniu pasirinkite vaizduojamą duomenų lauką ir, jei reikia, nurodykite x reikšmių lauką.
4. Jei norite sukurti diagramos legendą, pažymėkite legendos žymimąjį langelį.
5. Jei norite, pakeiskite juostų stilių ir spalvą.
6. Spustelėkite *Next* (toliau).
7. Pažymėkite visų objektų arba atrinktų objektų akutę.
8. Įveskite diagramos pavadinimą.
9. Ašių savybių (*Axis properties*) skyreliuose nustatykite ašių matomumą ir pavadinimus.
10. Spustelėkite *Finish* (baigti).



1.21 pav. 2 diagramų vedlio veiksmas (iš 2)

Dalies klausimai savarankiškam darbui

1. Kuo rastrinis duomenų modelis pranašesnis už vektorinį?
2. Kokie duomenų tipai (tekstas (*text*), ilgasis sveikasis skaičius (*long integer*) ar kt.) geriausiai tinka šiems žemės sklypų elementų klasės atributams saugoti:
 1. SklypoSavininkas (žemės savininko vardas ir pavardė)
 2. SklypoPlotas (sklypo plotas kvadratiniais metrais)
 3. PirkimoData (aiškus savaime)
 4. StatSkaičius (statinių skaičius sklype)
3. Kokios matavimų skalės (nominalinės, ranginės, intervalinės, santykinės ar ciklinės) pavyzdys yra žemėnaudos atributas (žemdirbystė, miškas, miestas ir pan.)?
4. Kuo skiriasi *ArcMap* jungimo (*Join*) ir siejimo (*Relate*) operacijos?
5. Kokia užklausa atrinktų visus objektus iš elementų klasės ŽemėsSklypai, kurių atributų lauko *Žemėnauda* reikšmė yra Žemdirbystė, o *PlotasHa* (plotas hektarais) didesnis negu 10 hektarų?

Literatūra

- Chang, K.T. *Introduction to Geographic Information Systems*. McGraw-Hill, 2006.
- Chrisman, N. *Exploring Geographic Information Systems*. John Wiley and Sons, 1997.
- Heywood, I., Cornelius, S. and Carver, S. *An Introduction to Geographical Information Systems*. Pearson Education, 2002,

2. Erdvinės analizės pagrindai

GIS erdvinė analizė leidžia paversti duomenis informacija ir, manipuliuojant esamais erdviniais elementais bei jų atributais, kurti naujus duomenis (išvestinius duomenų rinkinius). GIS paketai turi daugybę analizės funkcijų, galinčių manipuluoti ir rastriniais, ir vektoriniais duomenimis. Šias funkcijas galima laikyti savotišku erdvinės analizės įrankių rinkiniu – ir iš tiesų kelios GIS aplikacijos geografinio apdorojimo funkcijoms aprašyti naudoja šią „įrankinės“ analogiją.

Atlikdamas įvairias užduotis, GIS naudotojas paprastai naudojasi keliais iš šių analizės įrankių. Pavyzdžiui, imkime paprastą analizės užduotį – nustatyti žemės ūkio veiklos mastą 500 m atstumu nuo upių, tarkim, kiekybiškai įvertinant pakrančių zonų trikdžius arba vandens kokybės suprastėjimą. Kad atsakytume į šį klausimą, turime sukurti 500 m upių buferius, perdengti juos žemdirbystės tipo žemėnaudos poligonais ir apskaičiuoti gautos sankirtos plotą. Šios dalies tikslas – išvardyti ir apibrėžti daugumą įprastinių analizės įrankių, esančių daugumoje komercinių GIS aplikacijų, kuriuos kombinuojant galima spręsti konkrečius analizės uždavinius.

Organizaciniais sumetimais šias funkcijas suskirstėme į grupes ir nagrinėsime atskirose temose:

- Vektorinės analizės metodai
- Rastrinės analizės metodai
- Duomenų apibendrinimas

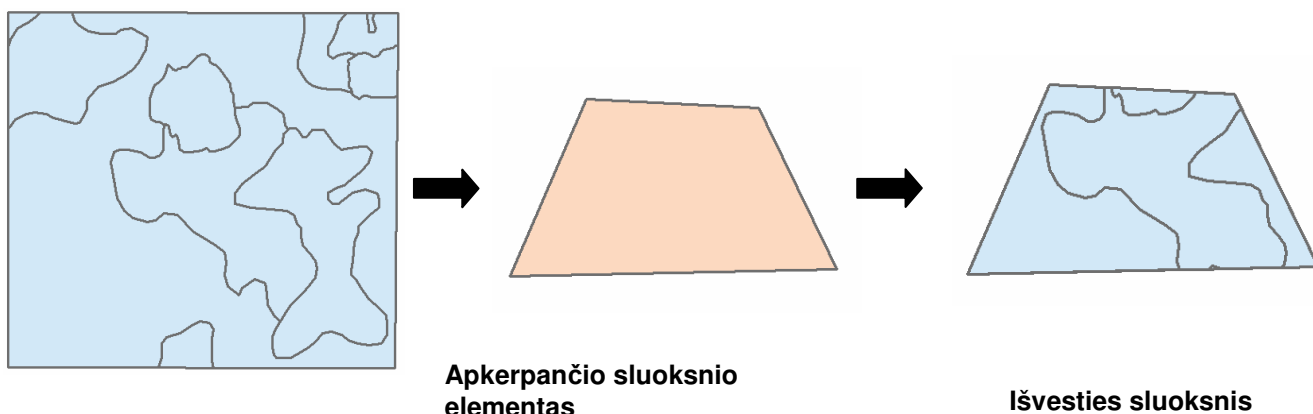
2.1 Vektorinės analizės metodai

2.1.1 Išskyrimas

Duomenų išskyrimas (*extraction*) yra efektyvus būdas atskirti konkrečias sritis tolesniam apdorojimui ar duomenų analizei. Panašiai kaip užklausos ir duomenų išrinkimo funkcijos, išskyrimo funkcijos tinka duomenų rinkinių dydžiui sumažinti ir (ar) sudėtingesnei interpretacijai palengvinti. Užklausos (*queries*) ir išrankos (*selection*) taip pat tinka duomenų rinkinio dalims atskirti, bet išskyrimo technikos skiriasi tuo, kad šios duomenų dalys atskiriamos visam laikui – sukuriama nauji duomenų sluoksniai. GIS programinės įrangos paketai turi visą komplektą įrankių duomenims išskirti – naudingiausi yra apkirpimas, išrinkimas, skaldymas ir šalinimas.

Apkirpimas

Apkirpimo (*Clip*) funkcija veikia panašiai kaip formelė sausainiams pjauti – galima sukirsti du elementų sluoksnius, išskiriant vieno duomenų rinkinio (įvesties sluoksnio) dalį pagal kito duomenų rinkinio (apkerpančio sluoksnio) erdvinę aprėptį. Apkirpimo funkcija sukuria naują (išvesties) duomenų sluoksnį, kurį sudaro įvesties sluoksnio elementai, patenkantys į apkerpančio sluoksnio aprėptį (2.1 pav.).



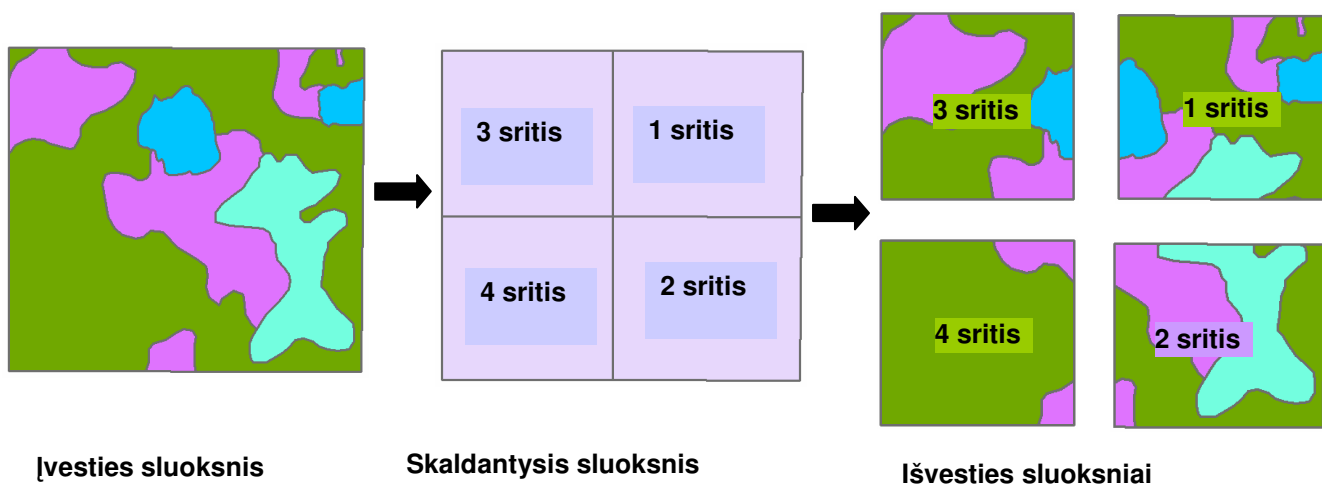
2.1 pav. Apkirpimo pavyzdys

Apkirpimas tinka tada, kai reikia iš esamų duomenų sluoksnių sukurti vieną jų poaibį, turintį bendras ribas, pavyzdžiui, tiriamosios srities arba administracines – rajono, apskrities, valstybės, savivaldybės ar pan. Pavyzdžiui, miesto projektuotojas gali norėti panagrinėti gatvių tinklo sluoksnį – bet tik tas gatves, kurios patenka į tam tikros savivaldybės ribas. Apkirpimas tiktų į savivaldybės ribas patenkantiems gatvių elementams išskirti visam laikui.

Apkerpamame įvesties sluoksnyje gali būti taškų, linijų ir poligonų elementų; bet kadangi reikia išskirti tam tikrą plotą, apkerpantis sluoksnis turi būti poligonas. Išvesties sluoksnio lentelėje išlieka laukų pavadinimai ir elementų atributai (t. y. jie yra tokie patys, kaip ir įvesties sluoksnio lentelėje). Galimos šios taisyklės išimtys – ploto, ilgio ir perimetro laukai, kurie (priklauso nuo programinės įrangos) gali būti arba nebūti automatiškai perskaičiuoti. Visų elementų, susikertančių su apkirpimo riba, atributų reikšmės gali tekti atnaujinti – kad atitiktų dėl kirpimo įvykusius pokyčius.

Skaldymas

Skaldymo (*Split*) funkcija dalina įvesties sluoksnį į du ar daugiau nepriklausomų sluoksnių pagal geografiškai atitinkančius skaldančiojo sluoksnio elementus. Panašiai kaip apkirpimo funkcijos atveju, įvesties sluoksnyje gali būti taškiniai, linijiniai arba poligoniniai elementai, bet skaldantysis sluoksnis turi būti poligonas, apibrėžiantis erdvinę analizės aprėptį. Įvesties sluoksnio elementai skaldomi ties skaldančiojo sluoksnio elementų ribomis, kaip pavaizduota 2.2 pav.

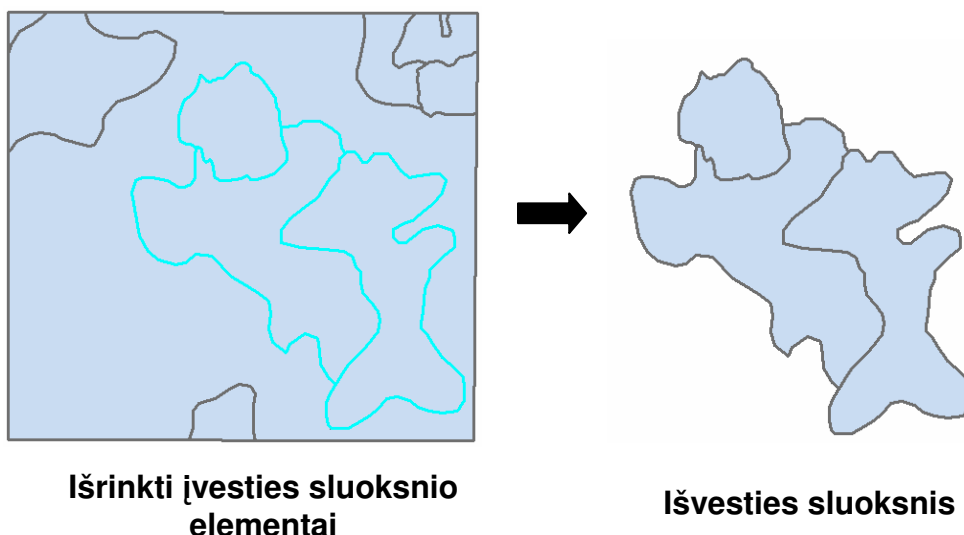


2.2 pav. Skaldymo pavyzdys

Iš principo skaldymas yra kelios apkirpimo operacijos, nurodytos skaldančiojo sluoksnio ribomis. Išvesties sluoksnių skaičių lemia skaldančiųjų elementų skaičius (t. y., jei skaldančiojo sluoksnio faile yra keturi poligonai, įvesties failas bus suskaldytas į keturis atskirus failus). Kiekvienam gautam išvesties sluoksniui bus suteiktas unikalus pavadinimas – pasirinkto skaldomojo sluoksnio atributo lauko reikšmė. Kaip ir apkerpant, skaldant išvesties sluoksniuose išlieka įvesties sluoksnio lentelės laukų pavadinimai ir atributai. Skaldymo funkcija tinka tada, kai didelės srities duomenis reikia padalinti į administracinius vienetus: pavyzdžiui, miesto zonų duomenis galima gauti padalinius pagal savivaldybių ribas, o valstybės žemėlapių rinkinį galima gauti padalinius topografinius duomenis nustatyto dydžio tinkleliu.

Išrinkimas

Išrinkimo (*Select*) įrankis sukuria naują sluoksnį iš elementų, išskirtų iš įvesties sluoksnio. Tai atliekama vykdant naudotojo sukurta užklausa, išrenkančią tam tikrą duomenų poaibį; išrinkti elementai visam laikui kopijuojami į naują duomenų sluoksnį (2.3 pav.).



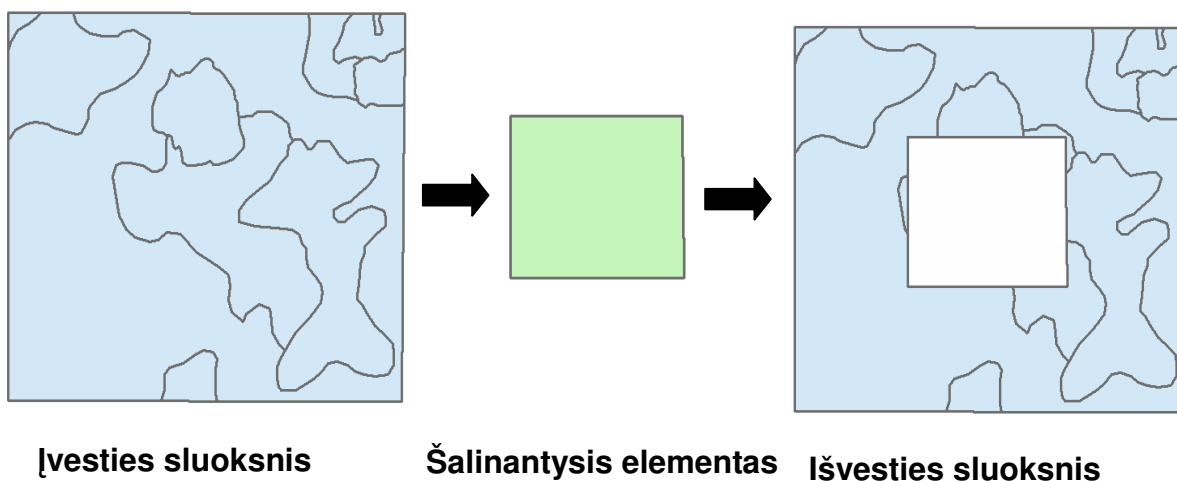
2.3 pav. Išrinkimo pavyzdys

Tęsdami ankstesnį pavyzdį, įsivaizduokime, kad miesto projektuotojas nori matyti tik dviejų juostų gatves tam tikroje savivaldybėje. Šiuo atveju jis įvykdytų išrinkimo užklausą ir išskirtų norimus elementus į naują sluoksnį.

Šalinimas (*Erase*)

Šalinant sukuriamas naujas išvesties sluoksnis be įvesties sluoksnio elementų, patenkančių į šalinantįjį (perdangos) sluoksnį (2.4 pav.). Įvesties sluoksnyje gali būti taškų, linijų ir poligonų elementų; bet kadangi reikia išskirti tam tikrą plotą, šalinantysis sluoksnis turi būti poligonas. Gauto išvesties sluoksnio elementų geometrijos tipas bus tas pat, kaip atitinkamų įvesties sluoksnio elementų. Šalinimo funkcijos naudojimo pavyzdžiai:

- žemėlapių maketui šalinimo funkcija galima sukurti tokį trafaretą, kad būtų rodomi tik į tam tikrą sritį (pvz., tyrimų srities ribas) patenkantys elementai;
- atliekant tinkamumo analizę, šalinimą galima naudoti tinkamumo taisyklėms taikyti. Pavyzdžiui, jei tinkamos vietos privalo būti ne arčiau kaip 200 metrų nuo pelkių, tada pelkėms galima sukurti 200 metrų pločio buferį ir buferio poligoną naudoti kaip šalinantįjį sluoksnį, į kurio ribas patenkančios vietos pašalinamos ir nebeagrinėjamos.



2.4 pav. Šalinimo pavyzdys

2.1.2 Perdanga (Overlay)

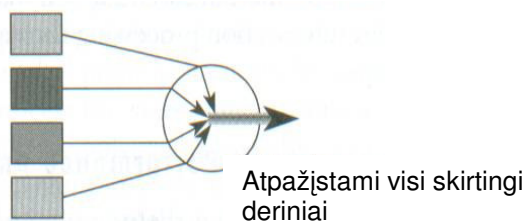
Viena pagrindinių GIS analizės funkcijų – apjungti du ar daugiau duomenų šaltinių, siekiant atsiskleisti jų tarpusavio ryšys (-iai). Vienas iš būdų informacijai apjungti yra perdanga – ši funkcija sujungia dviejų ar daugiau įvesties sluoksnių erdvinius ir atributų duomenis ir sukuria naują išvesties sluoksnį. Naujojo sluoksnio erdvinė forma priklauso nuo geometrinės įvesties bei perdangos sluoksnio elementų sankirtos. Perdengiant du ar daugiau sluoksnių, paprastai gaunamas sudėtingesnis išvesties sluoksnis – poligonų, sankirtų ir (ar) linijų atkarpų jame daugiau negu įvesties sluoksniuose.

Visų naujojo išvesties sluoksnio elementų atributai yra įvesties sluoksnio atributų kombinacijos. Perdangos funkcijos susiejamos su geometrinėmis (arba „fizinėmis“) duomenų sluoksnių perdangomis ir įgyvendinamos tam tikromis matematinėmis – aritmetinėmis ir loginėmis – operacijomis. Aritmetinės operacijos yra sudėtis, atimtis, dalyba ir daugyba ir kitos. Loginės operacijos skirtos nustatyti, kuriose vietose tenkinamos tam tikros sąlygos; jų operatoriai yra, pavyzdžiui, IR, ARBA, > ir <.

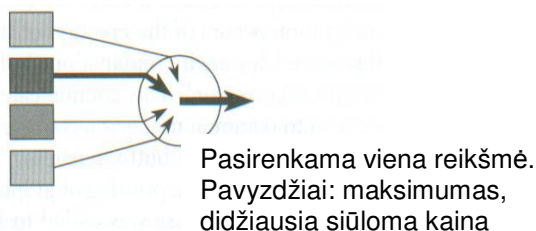
Vėliau šioje dalyje išsiaiškinsime, kad vektorinių duomenų perdangos metodai skiriasi nuo rastrinių duomenų perdangos metodų. Vektoriniai metodai duomenų neapibendrina, bet jiems reikia intensyvesnių skaičiavimų, todėl jie geriau tinka mažesniems arba retesniems duomenų rinkiniams. Rastrinės analizės metodai duomenis apibendrina pagal didžiausios įvesties sluoksnių gardelės dydį. Tačiau rastriniai skaičiavimai paprastai greitesni ir paprastesni.

Atliekant perdangos analizę, sluoksnių atributus galima jungti pagal keturias pagrindines taisykles, pavaizduotas 2.5 pav. Nors šios taisyklės aptariamos kartu su vektorinės analizės metodais, jos lygiai taip pat tinka ir rastrinės perdangos metodams.

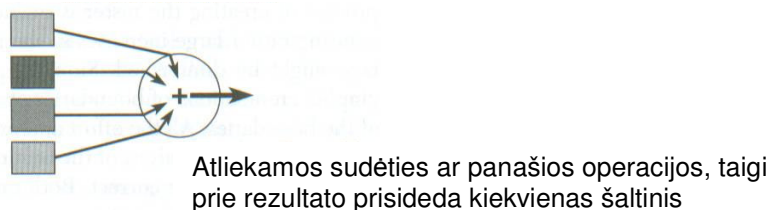
- Sąrašo taisyklė: išvesties sluoksnyje išsaugomas kiekvienas atributas ir atpažįstami visi unikalūs deriniai. Pavyzdžiui, perdengiami dirvožemio, augmenijos ir kritulių sluoksniai ir gaunamas išvesties sluoksnis, kuriame kiekvienam galimam deriniui priskiriamas unikalus poligonas.



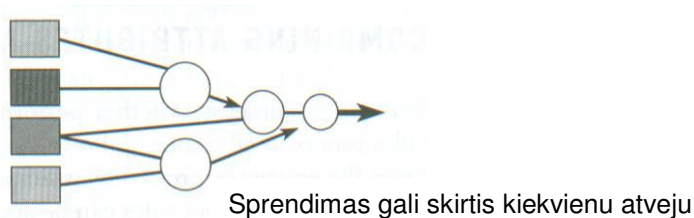
- Dominavimo taisyklė: laimi viena reikšmė, t. y. priskiriama tik viena – dominuojanti (pvz., didžiausia) reikšmė. Pavyzdžiui, perdengiant kelis jautrumo sluoksnius, išvesties poligonams priskiriamos didžiausios jautrumo reikšmės.



- Įnašo taisyklė: kiekviena atributo vertė (kiekvienas sluoksnis) prisideda prie rezultato. Pavyzdžiui, aplinkos jautrumą galima skaičiuoti sudedant laukinių gyvūnų buveinių jautrumo, pakrantės zonų jautrumo, nuolydžio ir žmogaus trikdžių artumo įvesties sluoksnių reikšmes.



- Sąveikos taisyklė: rezultatas priklauso nuo reikšmių poros (t. y. kiekvienu atveju sprendimas gali skirtis).



2.5 pav. Perdangos taisyklės

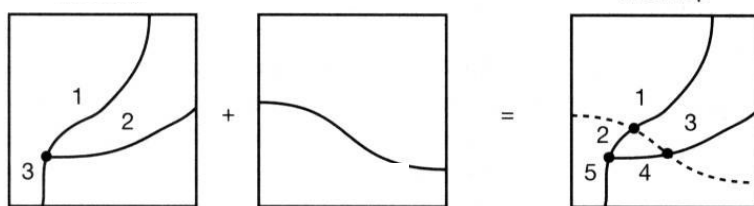
Vektorinės perdangos skirstomos į tris pagrindinius tipus:

- **taškai poligonuose** – taškiniams elementams, kurių vieta erdvėje ir atributai išlieka išvesties sluoksnyje, priskiriami ir atributai poligono, į kurį jie patenka. Šio tipo perdanga gali susieti, pavyzdžiui, meteorologines stotis (stočių taškinį sluoksnį) ir augmenijos tipus (miško poligonų sluoksnį) (2.6 pav.). Išvesties sluoksnis bus meteorologinių stočių taškų failas su papildomais augmenijos tipo atributais;



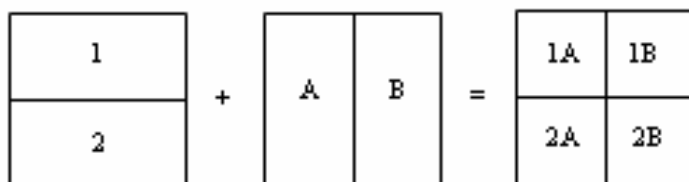
2.6 pav. Taškų perdengimo poligonais pavyzdys

- **linijos poligonuose** – linijiniams elementams, kurių vieta erdvėje ir atributai išlieka išvesties sluoksnyje, priskiriami ir atributai poligono, į kurį jie patenka. Šio tipo perdanga galima kiekvienai kelio linijai (kelių linijinis sluoksnis) nurodyti augmenijos tipus (gautus iš miško poligonų sluoksnio) (2.7 pav.). Kadangi viena linija gali persidengti su kelių tipų poligonais, išvesties sluoksnyje (kelių linijiniame žemėlapyje) greičiausiai bus daugiau linijų segmentų, negu įvesties sluoksnyje.



2.7 pav. Linijų perdengimo poligonais pavyzdys

- **poligonai ant poligonų** – iš įvesties ir perdangos sluoksnio gaunamas naujas poligonų rinkinys, kurio poligonai išsaugo abiejų sluoksnių atributus (8 pav.). Šio tipo perdanga tinka, pavyzdžiui, ieškant ryšio tarp nuolydžio ir lavinų kanalų. Poligonų perdengimas poligonais yra plačiausiai naudojamas vektorinės perdangos metodas.



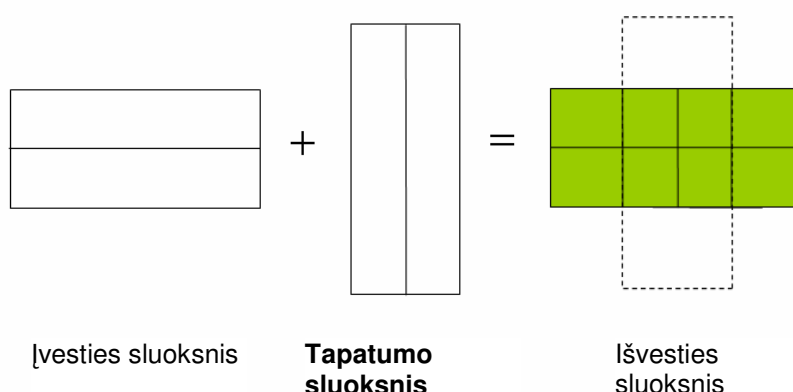
2.8 pav. Poligonų perdengimo poligonais pavyzdys

Perdengiant poligonus poligonais, poligonų rinkinius galima jungti keliais būdais:

- tapatumo,
- sankirtos,
- simetrinio skirtumo,
- sąjungos,
- atnaujinimo.

Tapatumas

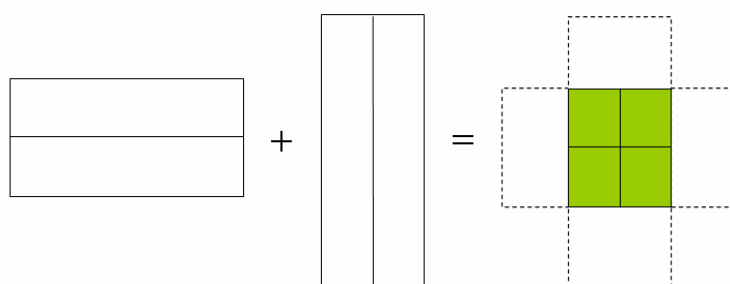
Tapatumas (*Identity*) – tai perdangos funkcija, iš įvesties sluoksnio sukurianti tos pačios erdvinės aprėpties sluoksnį (2.9 pav.). Visi įvesties sluoksnio elementai ir jų atributai išlieka. Ši operacija išsaugo ir į įvesties sluoksnio aprėptį patenkančios perdangos (arba tapatumo) sluoksnio dalies geometriją bei atributus. Įvesties sluoksnyje gali būti taškų, linijų ir poligonų, o tapatumo sluoksnio (dengiančiojo) elementai turi būti poligonai. Tapatumo funkciją galėtume naudoti, jei, pavyzdžiui, norėtume nustatyti kelius, esančius ne aukščiau negu 1 000 metrų virš jūros lygio. Įvesties sluoksnis būtų kelių sluoksnis, tapatumo sluoksnis – žemesnės negu 1 000 metrų sritis vaizduojantis poligonų sluoksnis. Gautame išvesties sluoksnyje žemiau negu 1 000 m esantys keliai būtų pažymėti aukščio atributu. Taip pat galėtume daryti išvadą, kad aukščio atributo neturintys keliai yra virš 1 000 metrų izolinijos.



2.9 pav. Tapatumo pavyzdys

Sankirta

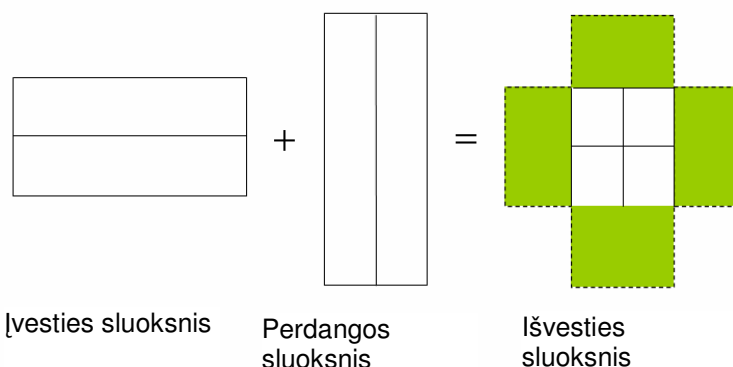
Sankirtos (*Intersect*) funkcija sukurtame išvesties sluoksnyje išlieka tik tie elementai (arba elementų dalys), kurie yra abiejuose įvesties sluoksniuose (2.10 pav.). Visi išvesties sluoksnio elementai turi abiejų įvesties sluoksnių atributus. Įvesties sluoksnių geometrija gali skirtis (taškai, linijos arba poligonai), bet paprastai įvedami poligonų sluoksniai. Išvesties sluoksnio geometrija supaprastinama iki žemiausio bendro įvesties sluoksnių lygmens – perdengiant taškinį sluoksnį poligonais, bus gautas taškų sluoksnis, ir taškai turės atributus poligonų, į kuriuos jie patenka. Sankirtos funkciją galėtume naudoti, jei, pavyzdžiui, norėtume išskirti kelius, esančius žemiau negu 1 000 metrų virš jūros lygio. Įvesties sluoksniai būtų kelių sluoksnis ir žemesnės negu 1 000 metrų sritis vaizduojantis poligonų sluoksnis. Gautame išvesties sluoksnyje būtų kelių poaibis – tik žemiau negu 1 000 m esantys keliai.



2.10 pav. Sankirtos pavyzdys

Simetrinis skirtumas

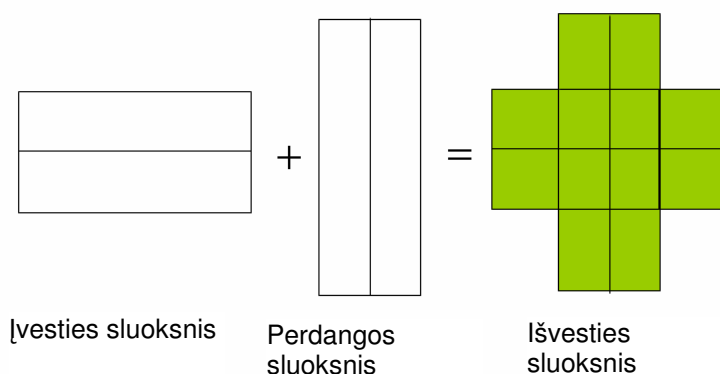
Simetrinio skirtumo (*symmetrical difference*) įrankis sukuria išvesties sluoksnį, kuriame išsaugomi nebendri įvesties sluoksnių elementai (arba jų dalys) – nepersidengiančios įvesties sluoksnių dalys (2.11 pav.). Šią operaciją galima atlikti tik su poligoniniais sluoksniais. Simetrinio skirtumo funkciją naudotume, pavyzdžiui, persidengiančioms įvesties sluoksnių sritims pašalinti – tarkim, jei turite poveikio zonos poligoną aplink chemijos gamyklą, šia funkcija galite nustatyti vietas, kurių neturėtų veikti gamyklos cheminė tarša.



2.11 pav. Simetrinio skirtumo pavyzdys

Sajunga

Sajungos (*union*) funkcija sujungia ir išsaugo visus abiejų įvesties sluoksnių elementus ir atributus (2.12 pav.). Šią operaciją galima atlikti tik su poligoniniais sluoksniais. Minėtos chemijos gamyklos atveju sąjungos operacija nustatytų ir į gamyklos poveikio zoną patenkančias, ir į ją nepatenkančias vietas.

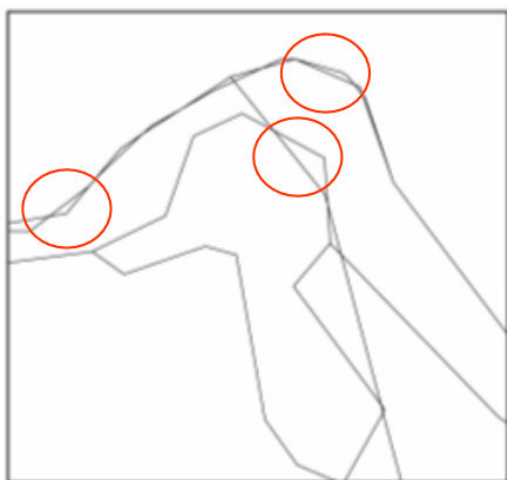


2.12 pav. Sąjungos pavyzdys

Atnaujinimas

Atnaujinimo (*update*) įrankis sukuria naują sluoksnį, pakeisdamas įvesties sluoksnio elementus bei atributus persidengiančiais atnaujinimo sluoksnio elementais ir jų atributais. Su atnaujinimo sluoksnio elementais nepersidengiančios įvesties sluoksnio dalys lieka nepakeistos, ir į išvesties sluoksnį įrašomos tokios, kokios buvo. Ši operacija tiktų atnaujinti žemėnaudos sluoksnį naujesniu miško kirtimų sluoksniu. Pasenusios žemėnaudos informacijos sritys, pavyzdžiui, neseniai iškirstos, būtų pašalintos ir vietoj jų būtų įrašyti kirtimų poligonai. Šią operaciją reikia atlikti atsargiai, nes dėl sluoksnių skirtumų (skirtingi duomenų šaltiniai, kita kokybė ir kitas mastelis) gali nesutapti ribos ir (arba) išvesties sluoksnyje atsirasti atplaišų.

Atplaišos (*slivers*) išvesties sluoksniuose yra viena iš būdingų klaidų, atsirandančių perdengiant poligonų sluoksnius. Tai labai maži poligonai ties bendromis perdengiamų sluoksnių ribomis (2.13 pav.). Atliekant perdangos analizę, ši problema gali atsirasti dėl skirtingo perdengiamų sluoksnių mastelio. Atplaišų gali atsirasti ir dėl įskaitmeninimo klaidų, netikslaus orientavimo geografinėje erdvėje arba duomenų eksporto. Padėti operatoriui pašalinti po poligonų perdangos operacijų atsiradusias atplaišas skirta likvidavimo funkcija (*Eliminate*).



2.13 pav. Atplaišų pavyzdys

Likvidavimas

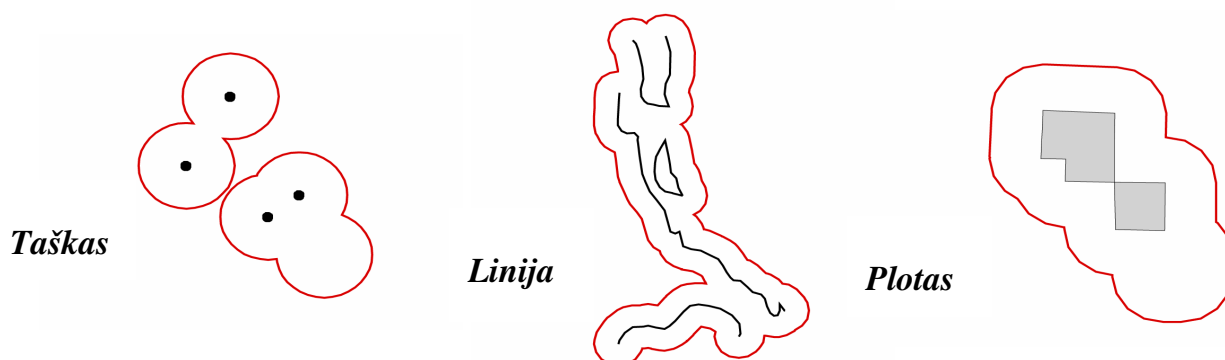
Likvidavimo komandą galima naudoti po perdengimo arba buferių kūrimo (žr. 2.1.3 skyrių) atsiradusioms atplaišoms pašalinti iš išvesties sluoksnio. Atplaišų poligonai prijungiami prie gretimų poligonų. Galima nurodyti, ar atplaišų poligonus prijungti prie didžiausio gretimo poligono, ar prie poligono su ilgiausia bendra riba. Paprastai likvidavimas vykdomas iškart po perdangos analizės – atliekama užklausa, atrenkanti mažesnius už nurodytą ribinį plotą (t. y. labai mažus poligonus – atplaišas), ir pagal šią išranką nustatomi poligonai, kuriuos reikia likviduoti. Likvidavimo funkciją reikia naudoti atsargiai, nes ji gali smarkiai pakeisti duomenų sluoksnį. Pavyzdžiui, perdengiant kelis sluoksnius atplaišos dažnai susidaro palei krantų linijas. Todėl įvykdžius likvidavimo komandą išvesties sluoksnyje gali pasikeisti kranto linija. Be to, išrenkant mažus poligonus reikia pasirinkti tinkamą ribinį plotą, nes jei jis bus per didelis, prie didelių gretimų poligonų gali būti prijungti ir nedideli „tikri“ poligonai.

2.1.3 Artimumas

Artimumas (*Proximity*) yra erdvinio ryšio sąvoka, atitinkanti geografinį elementų „artumą“. Ja remiantis galima išrinkti elementus, esančius ne toliau negu nurodytu atstumu nuo kitų elementų, arba sukurti naujus elementus, išplečiant esamų elementų aprėptį. Pavyzdžiui, galima rasti visus viešbučius, esančius ne toliau negu 10 kilometrų nuo Vilniaus centre esančios Katedros. Į panašius klausimus padeda atsakyti buferiai ir kita artimumo analizė.

Buferiai

Artimumo sąvoka pagrįstas buferių kūrimas – nurodyto pločio apimančios arba pašalinančios zonos suformavimas aplink esamus taškinius, linijinius ar poligoninius elementus. GIS programinė įranga aplink elementus brėžia liniją tol, kol susidaro uždaras poligonas, ir sukuria naują sluoksnį su gautais buferiais. Taškų buferiai yra apskriti, o linijų ir poligonų buferių forma priklauso nuo įvesties elementų geometrijos (1.14 pav.).



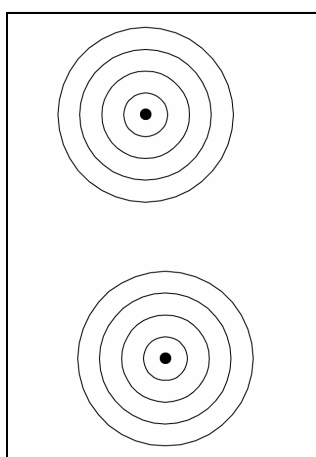
1.14 pav. Buferių tipai

Linijų buferius galima formuoti abiejose linijos pusėse, arba tik kairėje ar tik dešinėje. Poligonų buferių zonos vietą galima pasirinkti atsižvelgiant į dominančios srities padėtį – nurodyti, kad:

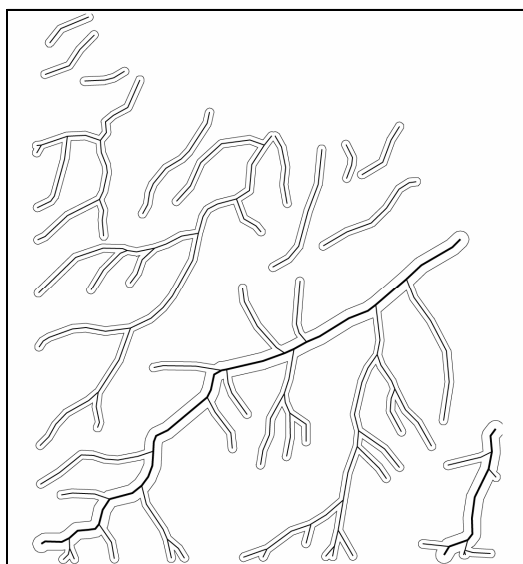
- buferis turi būti kuriamas tik poligono išorėje;
- į buferio zoną turi patekti sritis poligono išorėje ir visas poligono plotas;
- buferio zona turi būti kuriama ir poligono ribų viduje, ir išorėje.

Sukurtą buferį galima panaudoti keliais būdais. Tai gali būti naujas elementas – pavyzdžiui, pakrantės zona arba kelio poligonas, sukurti pagal upės ar kelio vidurio liniją. Taip pat jo pagrindu galima atlikti artimumo analizę – perdengti buferį kitu įvesties sluoksniu ir surasti su juo susikertančius elementus.

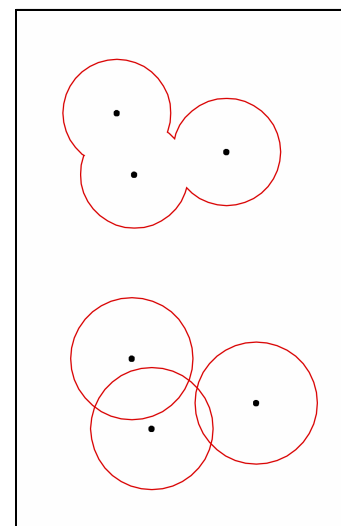
Buferio atstumas gali būti kintamas, taigi atskirų elementų buferiai gali būti skirtingo pločio; paprastai konkretaus buferio plotis priklauso nuo susijusioje atributų lentelėje nurodytos atstumo reikšmės. Pavyzdžiui, kuriant kintamo pločio pakrančių zonas pagal hidrologijos tinklą, didesnių upių buferius galima sukurti platesnius negu mažesnių upių ir upelių. Su kintamo buferio pločio sąvoka susijęs kelių buferių kūrimas aplink elementus. Aplink elementą galima sukurti kelis žiedus, vaizduojančius tam tikrą hierarchiją arba kintantį įtakos laipsnį. Ši funkcija tinka, pavyzdžiui, gamykloje nurodyti kintamo triukšmo lygio zonas aplink įrenginius. Kuriant buferius, taip pat galima sulieti gautus gretimus buferių poligonus (t. y. sujungti persidengiančias sritis) (2.15 pav.). Suliejimą išsamiau aptarsime 2.3.1 skyriuje.



Kelios zonos (žiedai)



Atstumu nurodytas kintamas buferio dydis



Sulietas (viršuje) ir nesulietas (apačioje) buferis

2.15 pav. Buferio kintamieji

Kad išvengtumėte klaidų, reikia žinoti konkretaus duomenų rinkinio matavimo vienetus (pvz., metrai arba kilometrai). Kadangi buferiai kuriami išrankai, prieš pradėdami juos kurti įsitikinkite, kad išrinkti būtent reikiami elementai.

Artumas (*Near*)

Artumo funkcija skaičiuoja atstumą tiesia linija nuo kiekvieno įvesties sluoksnio taško iki artimiausio kito sluoksnio (dar vadinamo „artimų elementų sluoksniu“ (*Near Feature layer*)) taško ar vietos linijoje. Apskaičiuotos atstumo reikšmės įrašomos į įvesties sluoksnio atributų lentelę. Šią funkciją galėtų naudoti gaisrininkai, norėdami nustatyti arčiausiai rajono mokyklų (įvesties sluoksnis) esančius vandens siurblius (artimų elementų sluoksnis).

Atstumas iki taško

Atstumo iki taško (*Point distance*) funkcija nustato atstumus nuo kiekvieno įvesties sluoksnio elemento iki visų nurodytu paieškos spinduliu esančių kito sluoksnio taškų. Rezultatai – unikalūs elementų identifikatoriai ir atstumų reikšmės – surašomi į išvesties lentelę. Šią funkciją minėtieji gaisrininkai galėtų naudoti tuo atveju, jei reikėtų nustatyti atstumus nuo kiekvienos mokyklos iki visų nurodytu spinduliu esančių vandens siurblių.

2.1.4 Statistika

Dažnis

Dažnio funkcija (*Frequency*) nurodytam vektorinio įvesties sluoksnio laukų rinkiniui sukuria lentelę su visais skirtingais kodais ir jų dažniu. Šia funkcija, pavyzdžiui, augmeniją tyrinėjantis mokslininkas galėtų nustatyti kiekvieno augalų tipo dažnį konkrečioje srityje.

Bendroji statistika

Bendrosios statistikos įrankis skaičiuoja vieną arba kelias skaitinių atributų lentelės laukų statistikas: sumą, vidurki, didžiausią reikšmę, mažiausią reikšmę, intervalą, standartinį nuokrypį, pirmąją reikšmę ir paskutiniąją reikšmę. Gautą lentelę galima išsaugoti įvairiais formatais, tarp jų ir *dbase* bei asmeninės geoduomenų bazės formatu.

2.2 Rastrinės analizės metodai

2.2.1 Analizės parinktys

Prieš atliekant rastrinio duomenų rinkinio analizę, reikia nustatyti analizės aplinkos parametrus, įskaitant analizės aprėptį ir išvesties gardelių dydį.

Analizės aprėptis

Jei analizuoti reikia tik tam tikrą didelio rastrinių duomenų rinkinio dalį, dominančią sritį galima nurodyti stačiakampiu – mažiausiomis ir didžiausiomis (x, y) koordinatėmis. Jei analizuojami keli rastriniai duomenų rinkiniai, analizės aprėptį galima nurodyti ir kitu būdu – tam tikra įvesties duomenų rinkinių kombinacija. Šiuo atveju sritį nurodo rastrų sąjungą arba sankirtą. Jei pasirenkama sąjungą, analizės aprėptį sudaro visų įvesties rastrų sritys. Jei pasirenkama sankirta, analizės sritį sudaro tik bendros visų įvesties duomenų rinkinių sritys (t. y. mažiausia galima sritis).

Trafaretai

Nurodžius trafaretą (*mask*), galima analizuoti tik pasirinktas gardeles, taigi trafaretai yra dar vienas analizės aprėpties nurodymo būdas. Atliekant analizę bus atsižvelgta tik į tas gardeles, kurias nurodo naudojamas trafaretas. Analizės trafaretus galima nurodyti dviem būdais:

- 1 atributu: galima išrinkti norimas rastro atributų lentelės eilutes, ir bus analizuojamos tik tam tikros rastro reikšmės. Pavyzdžiui, gali būti, kad lokių buveinės tiriantį ekologą domins tik virš nurodytos ribos esančios sritys;
- 2 sritimi: erdvinė analizės aprėptis nurodoma turimu vektoriniu (taškų, linijų ar poligonų) sluoksniu arba rastriniu duomenų rinkiniu. Bus analizuojamos tik į trafareto nurodytą sritį patenkančios gardelės. Pavyzdžiui, minėtasis ekologas gali norėti apriboti buveinių analizės sritį iki parko teritorijos – tada trafaretas turi būti parko aprėptį nurodantis poligonas.

Gardelės dydis

Išvesties rastro gardelės dydis labai priklauso nuo įvesties duomenų skiriamosios gebos arba gardelės dydžio. Bendra taisyklė tokia: išvesties rastro gardelė turi būti lygi arba didesnė už stambiausią įvesties rastrinių duomenų gardelę (vadinamą įvesties maksimumu). Tada išvesties rastro skiriama geba tikrai atitiks mažiausiai tikslų (grubiausią) duomenų rinkinį. Nurodžius mažesnę negu įvesties duomenų gardelės dydį, duomenų tikslumas nepadidės. Kai kurios kitos parinktys:

- **įvesties minimumas** – išvesties gardelės dydis lygus smulkiausiai įvesties duomenų gardelei. Su šia parinktimi būkite atsargūs – jau minėjome, kad smulkindami grubių įvesties duomenų gardeles tikslesnių, didesnės skiriamosios gebos rezultatų negausite;
- **nurodytas** – leidžia nurodyti konkretų išvesties gardelės dydį;
- **kaip sluoksnio** – reiškia, kad išvesties skiriama geba bus lygi įvesties rastro gardelės dydžiui.

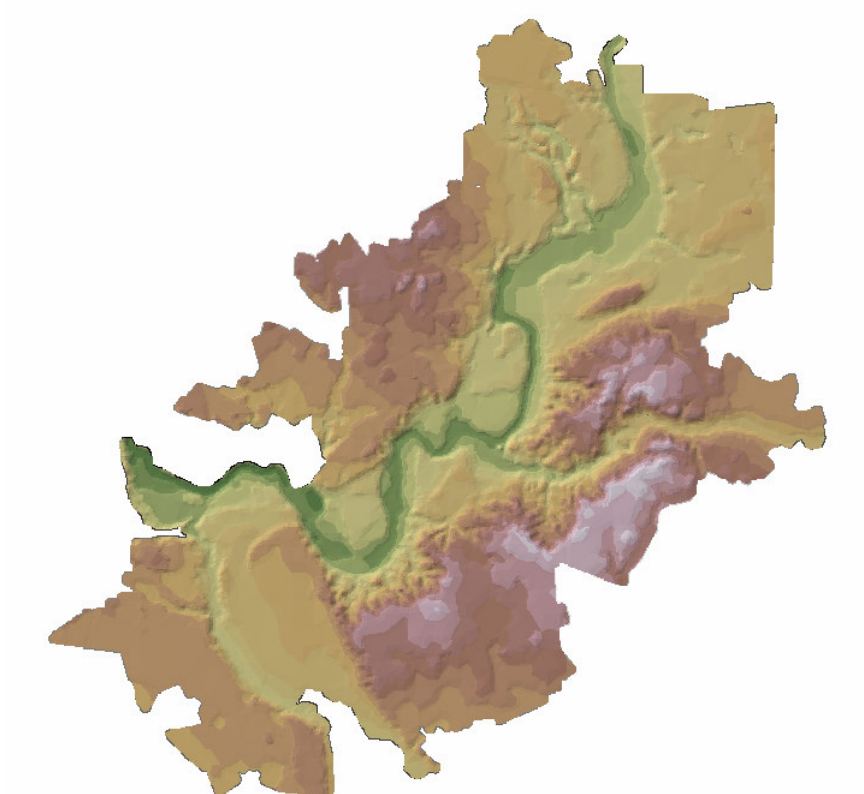
2.2.2 Paviršiaus analizė

GIS turi trimatės (3D) ir rastrinės analizės įrankių, kuriais galima tiksliai atvaizduoti, modeliuoti ir analizuoti reljefą. Toliau aprašyti paviršiaus analizės įrankiai leidžia nustatyti vietovės paviršiaus formas ir atskleisti dėsningumus, kurių galima ir neįžvelgti rastriniuose skaitmeniniuose aukščių modeliuose (DEM) (2.16 pav.) arba vektoriniuose netaisyklingų trikampių tinkluose (NTT) (2.17 pav.).

Šie reljefo analizės įrankiai išsamiai aptarti GII-04 kurse, o šiame kurse aprašyti tik trumpai, kad susidarytų visas vaizdas.



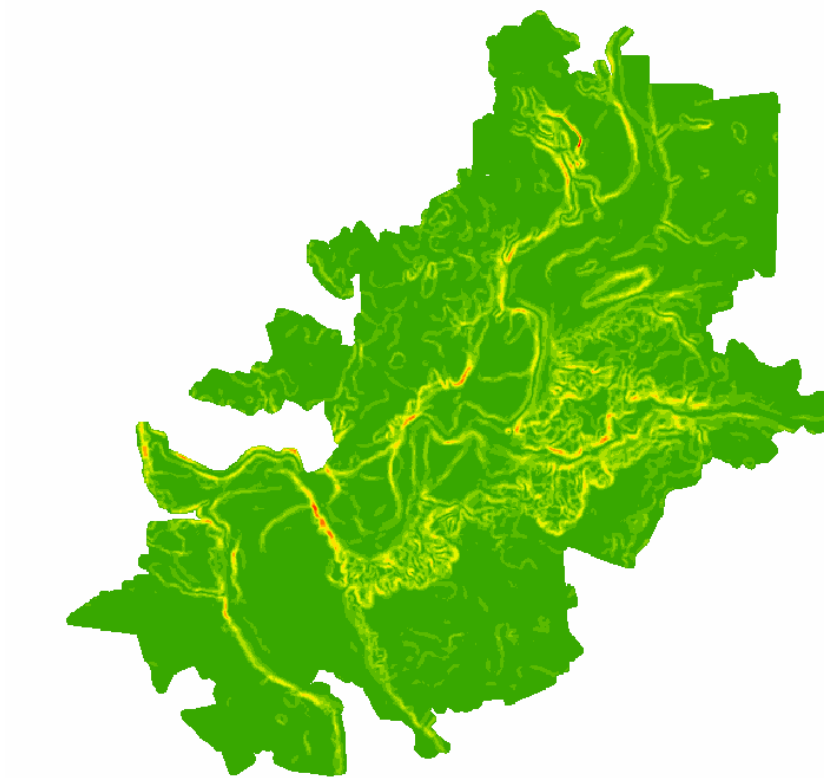
2.16 pav. Vilniaus skaitmeninis aukščių modelis



2.17 pav. Vilniaus reljefo NTT modelis

Nuolydis

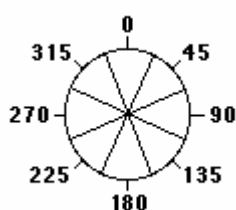
Nuolydis nurodo didžiausią aukščio pokytį konkrečioje paviršiaus vietoje. Nors jį galima išreikšti laipsniais arba procentais, abu šie dydžiai reiškia tą patį – aukščio padidėjimą ilgio vienetui (atstumui ant žemės paviršiaus). Išvesties rastras sudaromas atsižvelgiant į įvesties duomenų formatą – skaičiuojamas arba kiekvienos rastrinio DEM gardelės, arba kiekvieno NTT paviršiaus elemento nuolydis. Klasifikuodami rastrą pagal nuolydžio reikšmių grupes (pvz., $0-2^\circ$, $2-4^\circ$ ir pan.), galime atskirti nedidelio, vidutinio ir didelio nuolydžio sritis (2.18 pav.). Pavyzdžiui, matyti, kad statesnės sritys (geltonos ir raudonos) yra Vilnios ir Neries krantuose. Tokio pobūdžio nuolydžio informaciją galėtų susikurti ir analizuoti, pavyzdžiui, statybos inžinierius, norėdamas sumažinti nuošliaužų pavojų.



2.18 pav. Vilniaus nuolydžio žemėlapis

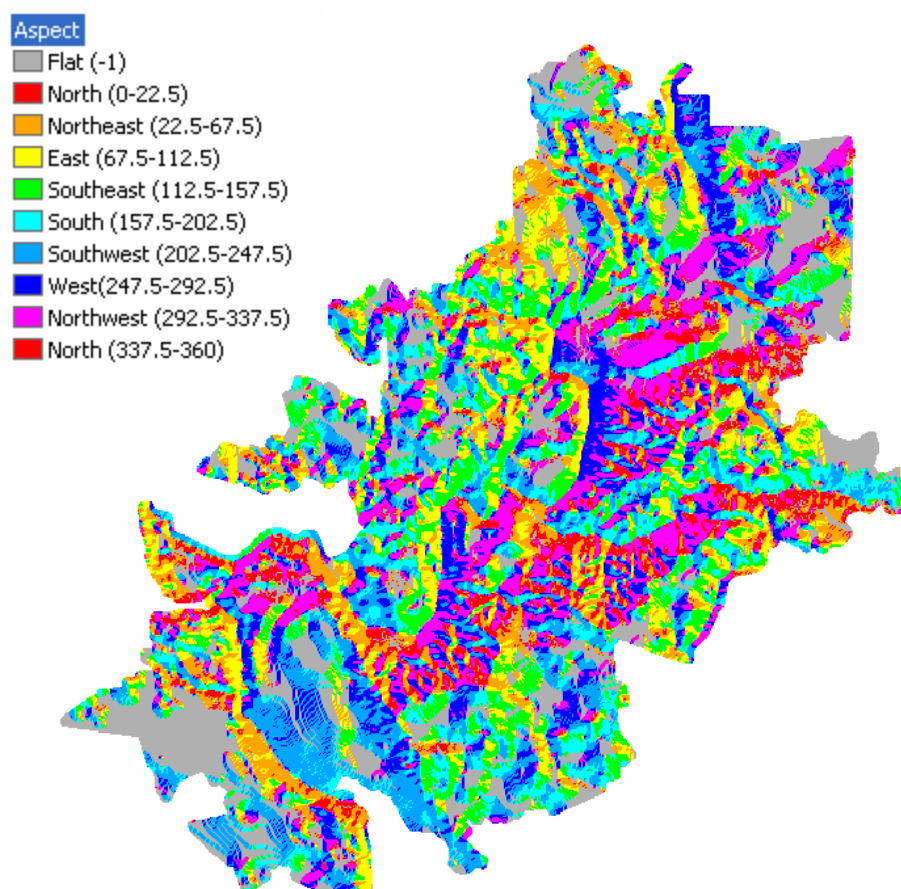
Ekspozicija (*Aspect*)

Ekspozicija apibrėžiama kaip didžiausio aukščio kitimo kryptis; iš esmės ji nurodo pagrindinę šlaito kryptį (pvz., į šiaurę, į pietus, į rytus, į vakarus). Ekspozicija matuojama pagal laikrodžio rodyklę nuo šiaurės (0°) ir išreiškiama laipsniais nuo 0° iki 359.9° (2.19 pav.).



2.19 pav. Ekspozicija laipsniais nuo šiaurės pagal laikrodžio rodyklę

Kaip ir nuolydžio rastro atveju, sudarant ekspozicijos išvesties rastrą skaičiuojama arba kiekvienos rastrinio DEM gardelės, arba kiekvieno NTT paviršiaus elemento ekspozicija – tai priklauso nuo įvesties duomenų rinkinio formato. Gardelėms, kurių nuolydis lygus nuliui, priskiriama ekspozicijos reikšmė „-1“. Iš 2.20 pav. matyti, kad ekspozicijos sluoksnis atskleidžia reljefo dėsninumus, kurių neįžvelgtume paprastame DEM ir net nuolydžio žemėlapyje.



Aspect	Orientacija
Flat (-1)	Horizontalus (-1)
North (0-22.5)	Šiaurė (0-22,5)
Northeast (22.5-67.5)	Šiaurės rytai (22,5-67,5)
East (67.5-112.5)	Rytai (67,5-112,5)
Southeast (112.5-157.5)	Pietryčiai (112,5-157,5)
South (157.5-202.5)	Pietūs (157,5-202,5)
Southwest (202.5-247.5)	Pietvakariai (202,5-247,5)
West (247.5-292.5)	Vakarai (247,5-292,5)
Northwest (292.5-337.5)	Šiaurės vakarai (292,5-337,5)
North (337.5-360)	Šiaurė (337,5-360)

2.20 pav. Vilniaus paviršiaus ekspozicijos žemėlapis

Tarkime, kad hidrologas nori nustatyti būsimas sniego dangos mėginių ėmimo vietas – gali būti, kad iš visų ekspozicijų jį domins tik šiauriniai ir šiaurės rytų šlaitai.

Kuriant nuolydžio arba ekspozicijos rastrus, kaip ir kiekvieno rastrinio duomenų rinkinio atveju reikia įvertinti skiriamąją gebą. Kaip taisyklė, išvesties rastro gardelė neturi būti mažesnė už įvesties rastro gardelę.

Izolinijos (*Contour*)

Izolinių įrankis sukuria izolinių – vienodo aukščio linijų (horizontalių) sluoksnį (2.21 pav.). Iš izolinių žemėlapių galima nustatyti vienodo aukščio sritis, o pagal izolinių padėtį ir tarpus tarp jų galima spręsti, kur yra statūs šlaitai, uolos, upių slėniai ir kalnų keteros. Stataus reljefo vietose izolinijos yra arti viena kitos, o lygumose – labiau nutolusios. Srityse aplink vandens objektus izolinių smailės nurodo priešingą upių tėkmei kryptį. Kuriant izolinijas, svarbu nustatyti tinkamą intervalą – aukščio skirtumą tarp gretimų izolinių. Didesnės skiriamosios gebos DEM arba NTT galima nurodyti mažesnį izolinių intervalą, negu stambesnių gardelių paviršiams.



2.21 pav. Vilniaus izolinių žemėlapis

Izolinių sluoksnį galima naudoti, pavyzdžiui, galimoms vilkų buveinėms (virš tam tikro aukščio) nustatyti.

Reljefo šešėliavimas

Reljefo šešėliavimo (*Hill shade*) įrankis iš aukščių rastro arba NTT sukuria reljefo rastrą su šešėliais. Ši funkcija gali labai gerai pavaizduoti vietovę ar reljefą: paviršiaus sluoksnis „apšviečiamas“, vaizduojami krentantys šešėliai ir taip sudaromas trimačio vietovės vaizdo įspūdis (2.22 pav.). Reljefo šešėliavime reikiami keturi parametrai:

- saulės azimutas – krentančios šviesos kryptis laipsniais nuo šiaurės pagal laikrodžio rodyklę (0° – 360°);
- Saulės altitudė – šviesos šaltinio kampas laipsniais virš horizonto (0° – 90°);
- paviršiaus nuolydis – DEM gardelės ar NTT paviršiaus elemento nuolydis;
- paviršiaus ekspozicija – DEM gardelės ar NTT paviršiaus elemento ekspozicija.

Visoms šešėlių rastro gardelėms priskiriamos apšviestumo reikšmės (nuo 0 – juodos iki 255 – baltos), kurios visos kartu sudaro trimačio reljefo įspūdį.



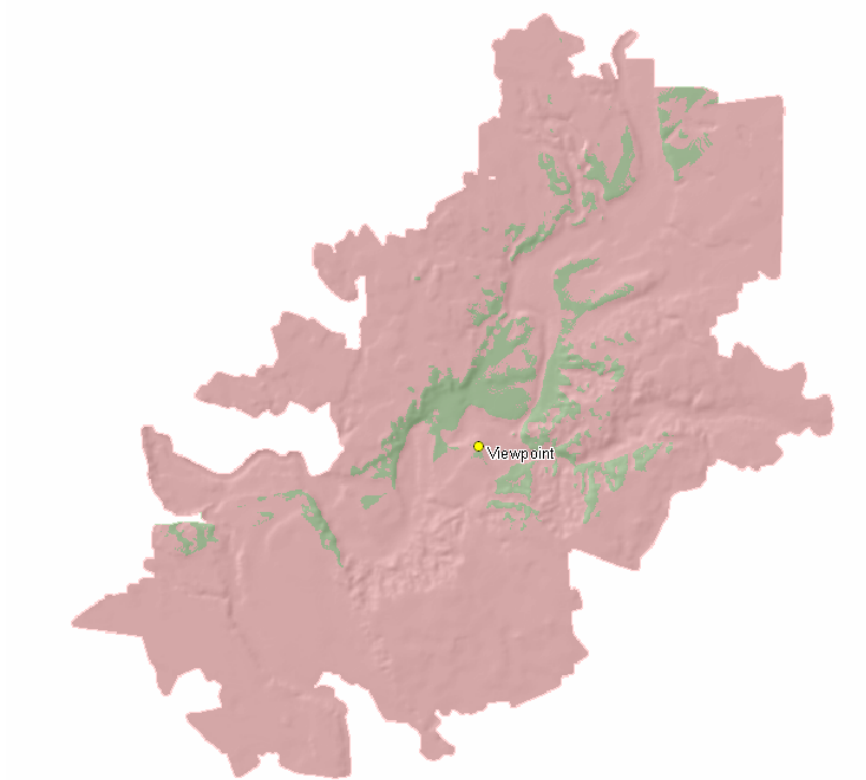
2.22 pav. Vilniaus šešėlių žemėlapis

Dažnai šešėliai naudojami kaip teminių žemėlapių fonas, perteikiantis reljefo įspūdį, bet neužgožiantis kitos, svarbesnės žemėlapių informacijos. „Pakišti“ šešėlių sluoksnius po pagrindiniais duomenimis galima nustatant poligoninių arba rastrinių sluoksnių skaidrumo parametą.

Matomumas

Matomumo (*Viewshed*) analizė atsako į klausimą, kurie elementai arba kurios paviršiaus sritys matomos iš vieno ar kito stebėjimo taško. Matomumo analizei reikia dviejų duomenų rinkinių: pirmasis – taškinių elementų sluoksnis su vienu ar daugiau stebėjimo taškų, antrasis – vietovę vaizduojantis DEM arba NTT paviršius.

Matomumo analizė pagrįsta „tiesioginio matomumo linijos“ sąvoka; ši linija jungia stebėjimo tašką ir stebimą objektą. Jei stebimą objektą užstoja koks nors kitas objektas (pvz., kalnas), stebimas objektas yra nematomas. Ir atvirkščiai, jei stebimo objekto niekas neužstoja, stebimas objektas yra matomas. Matomumo analizės atveju stebimi objektai yra visos įvesties paviršiaus sluoksnio gardelės arba paviršiaus elementai. Sudarant išvesties rastrą, sumuojami visi galimi tiesioginio matomumo linijų deriniai ir kiekviena gardelė klasifikuojama kaip „matoma“ arba „nematoma“ (2.23 pav.). Kad sumažintų miško kirtimo darbų poveikį kraštovaizdžiui, medienos ruošos įmonė pagal matomumo žemėlapi gali apriboti šiuos darbus iš gyvenviečių arba kelių matomose srityse.



2.23 pav. Matomumo analizės pavyzdys

Tam tikromis parinktimis galima valdyti matomumo analizės vykdymą ir apriboti arba pakeisti rezultatus. Aptarsime kelis pavyzdžius iš *ArcGIS* paketo. Šie kintamieji gali būti skaitiniai stebimojo sluoksnio atributų lentelės laukai, kurių reikšmės yra stebėjimo apribojimai:

- SPOT: z-reikšmė (aukštis), pakeičianti iš paviršiaus z-reikšmės gaunamą aukščio reikšmę;
- OFFSETA: nurodo, kokį vertikalų atstumą reikia pridėti prie stebėjimo taško z-reikšmės (pvz., 1,5 metro – vidutinį žmogaus ūgį);
- OFFSETB: nurodo, kokį vertikalų atstumą reikia pridėti prie kiekvienos stebimo paviršiaus gardelės z-reikšmės (pvz., vidutiniam augmenijos aukščiui įvertinti);
- AZIMUTH1: pirmoji iš dviejų reikšmių (laipsniais), ribojanti matomumo kampą iki tam tikros krypties;
- AZIMUTH2: antroji iš dviejų reikšmių (laipsniais), ribojanti matomumo kampą iki tam tikros krypties;
- VERT1: nurodo viršutinę vertikalaus matomumo kampo ribą;
- VERT2: nurodo apatinę vertikalaus matomumo kampo ribą;
- RADIUS1: nurodo spindulį, nuo kurio pradedama matomumo analizė (t. y. analizuojant neatsižvelgiama į arčiau negu RADIUS1 atstumu esančias rastro gardeles);
- RADIUS2: nurodo spindulį, kuriuo ribojama matomumo analizė (t. y. analizuojant neatsižvelgiama į toliau negu RADIUS2 atstumu esančias rastro gardeles).

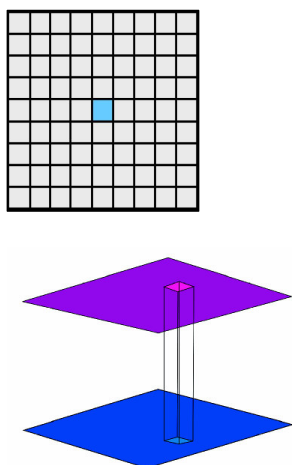
Bendrosios rastrų funkcijos (tinkamos ne tik specifiniams atvejams, pavyzdžiui, reljefui) skirstomos į tris kategorijas:

- vietinės funkcijos nagrinėja atskiras rastro gardeles;
- židinio, arba kaimynystės funkcijos nagrinėja židinio gardelę ir aplinkines arba ne toliau negu nurodytu atstumu esančias gardeles;
- zoninės funkcijos nagrinėja bet kokios formos vienodų reikšmių rastro gardelių rinkinius.

Šios funkcijos išsamiai aprašytos tolesniuose skyriuose.

2.2.3 Vietinės funkcijos ir vietinė statistika

Vietinės funkcijos vienu metu atlieka skaičiavimus su viena gardele, nepaisydamos gretimų gardelių reikšmių – aplinkinės gardelės apdorojamai gardelei įtakos neturi. Apskaičiavusi vieną gardelę, funkcija ima skaičiuoti kitą ir taip toliau, kol apdorojamos visos rastro (arba trafareto) gardelės (2.24 pav.).



2.24 pav. Vietinė funkcija

Vietinės operacijos naują rastrą gali sukurti iš vieno arba iš kelių įvesties rastrų.

Vieno rastro vietinės funkcijos

Vietinės funkcijos su kiekviena vieno įvesties sluoksnio gardele gali atlikti bet kokias aritmetines operacijas. Pavyzdžiui, jei reikia konvertuoti kritulių duomenis iš colių į milimetrus, galima padauginti kiekvieną gardelę iš 25,4 (2.25 pav.).

0.5	0.7	1
0.7	0.7	1
1	1	1

 $\times 25.4 =$

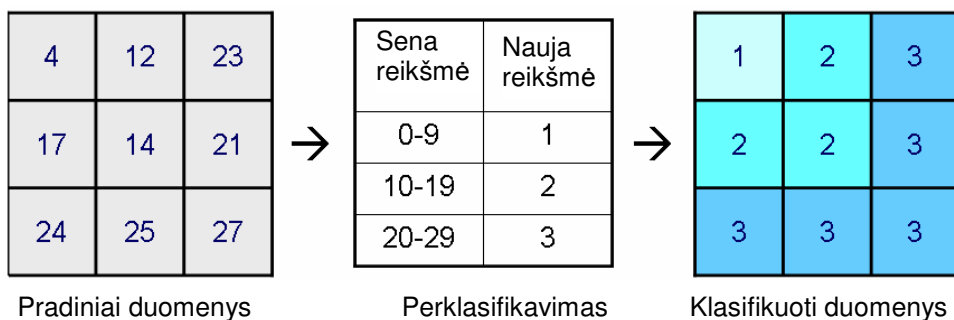
12.7	17.8	25.4
17.8	17.8	25.4
25.4	25.4	25.4

Įvesties sluoksnis

Išvesties sluoksnis

2.25 pav. Vietinė vieno sluoksnio funkcija

Perklasifikavimas yra vietinė funkcija, pakeičianti vienas įvesties rastro reikšmes kitomis ir sukurianti naują išvesties rastrą. Šia funkcija galima supaprastinti arba agreguoti (apjungti duomenis į klases) rastrinio duomenų rinkinio gardelių reikšmes (2.26 pav.).



Pradiniai duomenys

Perklasifikavimas

Klasifikuoti duomenys

2.26 pav. Perklasifikavimo pavyzdys

Perklasifikavimas dažnai naudojamas ir tada, kai gardelių reikšmes reikia pakeisti pagal naują informaciją arba „No Data“ reikšmes pakeisti tikromis reikšmėmis (kai duomenų rinkinyje yra tuščių vietų). Perklasifikavimas išsamiau nagrinėjamas 1.3.2 skyriuje.

Kelių rastrų vietinės funkcijos

Vietines funkcijas galima atlikti ir su keliais sluoksniais. Funkciniu požiūriu tai yra rastrų perdangos operacija, galinti jungti įvesties sluoksnių atributus įvairiais būdais, kuriuos jau aptarėme nagrinėdami vektorinę perdangą. 2.5 pav. pavaizduotas atributų jungimo taisyklės galima taikyti ir rastriniams sluoksniams. Pavyzdžiui, dominavimo taisyklė didžiausią reikšmę iš įvesties sluoksnių gali nustatyti maksimumo funkcija (*MAXIMUM*), įnašo taisyklė gardelių reikšmes gali jungti įvairiomis aritmetinėmis operacijomis, o sąveikos taisyklė gali priimti nuo gardelės reikšmės priklausančius sprendimus keliomis sąlygos (*IF..ELSE*) konstrukcijomis. Dauguma komercinių GIS aplikacijų turi įvairių priemonių rastrų gardelėms perdengti. *ArcGIS* programoje ši sąsaja vadinama rastrų skaičiuotuvu (*Raster Calculator*). 2.27 pav. pateiktas rastrų gardelių jungimo dviem aritmetinėmis operacijomis pavyzdys.

9	16	23
16	16	23
23	23	23

įvesties sluoksnis A

+

8	8	17
8	8	17
22	22	17

įvesties sluoksnis B

=

17	24	40
24	24	40
45	45	40

išvesties sluoksnis

9	16	23
16	16	23
23	23	23

įvesties sluoksnis A

/

8	8	17
8	8	17
22	22	17

įvesties sluoksnis B

=

1.13	2.00	1.35
2.00	2.00	1.35
1.05	1.05	1.35

išvesties sluoksnis

2.27 pav. Vietiniai veiksmai su keliais sluoksniais

Vietinė statistika

Kitas vietinių funkcijų pritaikymas yra vietinė (gardelių) statistika. Atliekant šią operaciją, dažnai vadinamą kompozicija (*compositing*), perdanga ar superpozicija (*superimposing*), skaičiuojama ir į išvesties rastrą įrašoma konkreti kelių rastrinių sluoksnių statistika. Kadangi vietinė statistika yra vietinė funkcija, ji lygina ir sumuoja tik atitinkamas įvesties rastrų gardeles (t. y. analizė atliekama po vieną gardelę).

Štai keli rastrinių duomenų rinkinio statistikų pavyzdžiai:

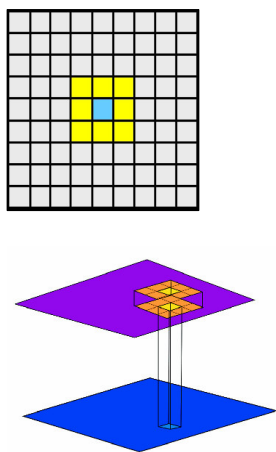
- maksimumas – nustato didžiausias atitinkamų įvesties rastrų gardelių reikšmes;
- minimumas – nustato mažiausias atitinkamų įvesties rastrų gardelių reikšmes;
- dauguma – nustato, kuri reikšmė atitinkamose įvesties rastrų gardelėse dažniausia;
- mažuma – nustato, kuri reikšmė atitinkamose įvesties rastrų gardelėse rečiausia;
- suma – skaičiuoja atitinkamų įvesties rastrų gardelių sumas;
- vidurkis – skaičiuoja vidutines atitinkamų įvesties rastrų gardelių reikšmes (vidurkius);
- mediana – skaičiuoja vidurines atitinkamų įvesties rastrų gardelių reikšmes (pusė reikšmių mažesnės, pusė – didesnės);
- standartinis nuokrypis – skaičiuoja atitinkamų įvesties rastrų gardelių standartinį nuokrypį;
- intervalas – nustato atitinkamų įvesties rastrų gardelių reikšmių intervalą (nuo mažiausios iki didžiausios reikšmės);
- įvairovė – nustato, kiek skirtingų reikšmių yra atitinkamose įvesties rastrų gardelėse.

Pavyzdžiui, nustatant didžiausias kelių rastrų reikšmes, kiekvienos išvesties rastro gardelės reikšmė nustatoma lyginant atitinkamas įvesties rastrų gardelių reikšmes. Galima, tarkime, iš penkių įvesties rastrų, vaizduojančių kritulių kiekius per penkis metus iš eilės, sudaryti vieną, vaizduojantį didžiausią metinį kritulių kiekį kiekvienoje gardelėje.

Loginėse funkcijose dažnai naudojamos ne tik aritmetinės, bet ir loginės (AND, OR, XOR, NOT ir pan.) bei palyginimo (pvz., >, <, =) operacijos.

2.2.4 Židinio funkcijos ir židinio statistika

Židinio (arba kaimynystės) funkcijų aprėptis platesnė negu vietinių ta prasme, kad nustatant išvesties rastro gardelės reikšmę atsižvelgiama ir į nustatytoje kaimynystėje esančių gardelių reikšmes. Dažniausiai naudojamos kaimynystės konfigūracijos yra stačiakampis, skritulys arba žiedas. 2.28 pav. mėlynas langelis vaizduoja židinio gardelę, o geltoni – 3x3 stačiakampės konfigūracijos gardelių kaimynystę (pati židinio gardelė į kaimynystę nebūtinai įtraukiama).



2.28 pav. Židinio funkcija

Panašiai kaip vykdant vietinę funkciją, funkcija ima gardeles po vieną (jos tampa židinio gardelėmis) tol, kol apdorojamos visos rastro (arba trafareto) gardelės. Skirtumas nuo vietinės kelių rastrų funkcijos tas, kad nustatydamas išvesties sluoksnio reikšmes židinio funkcija atsižvelgia ir į aplinkinių gardelių reikšmes.

Židinio statistika

Židinio funkcijos į išvesties gardeles gali įrašyti tas pačias statistikas, kaip ir vietinės operacijos. Pavyzdžiui, 2.29 pav. pavaizduota, kaip galima naudoti sumos funkciją, jei išvesties rastro židinio gardelėse turi būti įvertinti ir aplinkinių gardelių duomenys.

2	4	1	2	3
3	2	1	3	2
4	2	3	1	0
1	4			1
1	2	2	4	0

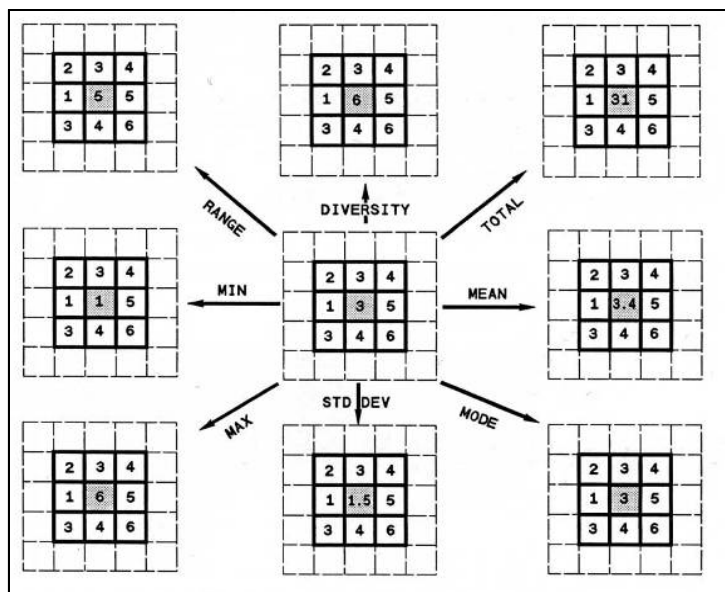
	22	19	16	

2.29 pav. Židinio sumos statistikos pavyzdys

Duomenų supaprastinimas

Kitas svarbus židinio statistikos funkcijų taikymas yra duomenų supaprastinimas (*simplification*), dar vadinamas erdviu filtravimu (*spatial filter*). Šio tipo operacijos tinka rastriniams duomenims apibendrinti ir skirtos sumažinti duomenų kintamumą gretimose įvesties sluoksnio gardelėse. Tipinis supaprastinimo variantas, vadinamasis judančio

vidurkio metodas, vidurkio (*Mean*) operatoriumi skaičiuoja vidutinę 3×3 arba 5×5 gardelių stačiakampio reikšmę; apskaičiuota vidutinė kaimynystės reikšmė įrašoma į šios kaimynystės židinio gardelę išvesties sluoksnyje. 2.30 pav. pavaizduotos tipinės rastrų erdvinio filtrų statistikos ir jų rezultatai.

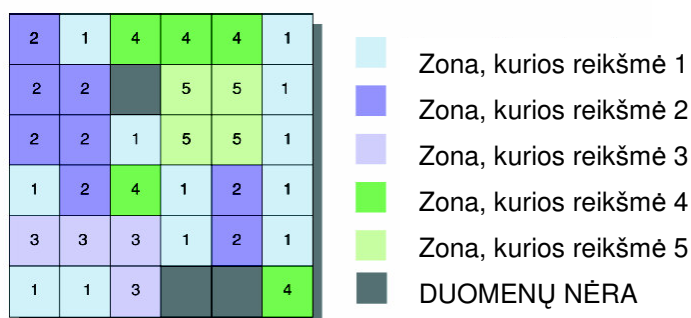


RANGE	INTERVALAS
DIVERSITY	VAIROVĖ
TOTAL	SUMA
MIN	MINIMUMAS
MEAN	VIDURKIS
MAX	MAKSIMUMAS
STD DEV	STD. NUOKRYPIS
MODE	MODA

2.30 pav. Tipiniai židinio statistikų erdviniai filtrai

„No Data“ reikšmės židinio statistikose

Vienas praktinis rastrinių duomenų analizės aspektas susijęs su tuščiomis gardelėmis. Šioms gardelėms galima priskirti reikšmę „duomenų nėra“ (*No Data*). Ši reikšmė rodo, kad gardelei neįmanoma priskirti reikšmės, nes nėra arba nepakanka informacijos (2.31 pav.).



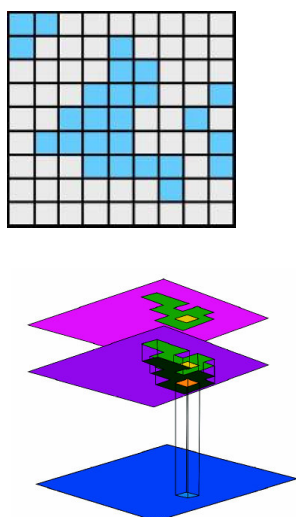
2.31 pav. Rastrinių duomenų „No Data“ reikšmės

Vykdamas vietines arba židinio funkcijas, gardeles su „No Data“ reikšmėmis galima apdoroti dvejopai:

- neatsižvelgiant į kitas įvesties gardeles, išvesties gardelei priskirti reikšmę „No Data“ (t. y., jei nors vienos kaimynystės gardelės reikšmė yra „No Data“, rezultatas bus „No Data“);
- gardelių su „No Data“ reikšmėmis nepaisyti ir atlikti skaičiavimus be jų (pavyzdžiui, nekreipti dėmesio į tuščią gardelę ir skaičiuoti didžiausią likusių gardelių reikšmę). Paprastai gardelių su „No Data“ reikšmėmis nepaisoma (pvz., skaičiuojama visų reikšmes turinčių gardelių suma).

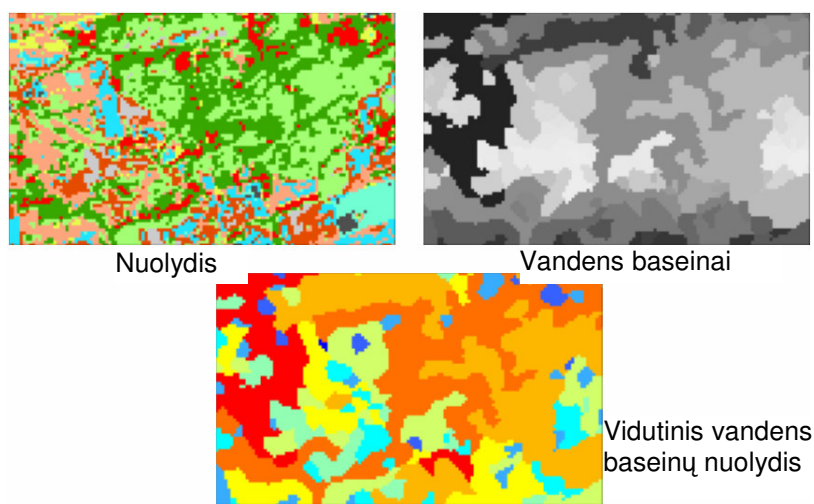
2.2.5 Zoninės funkcijos ir zoninė statistika

Zoninės funkcijos atlieka operacijas su vieno rastro gardelėmis, patenkančiomis į zonas, sudarytas iš vienodų kito rastro gardelių (t. y. su įvairios formos ir įvairaus dydžio kaimynystėmis). Zoną sudaro vienodas reikšmes turinčios gardelės. Zonos gali būti vientisos arba nevientisos (2.32 pav.). Vientisa zona sudaryta iš erdvėje besijungiančių gardelių, nevientisa – iš atskirų gardelių grupių.



2.32 pav. Rastrinio duomenų rinkinio zonos

Zonines operacijas galima atlikti su vienu arba dviem rastriniais sluoksniais (vienaime rastrė yra sumuojamos reikšmės, kitame nurodytos zonos). Jei įvedamas vienas rastras, zoninės statistikos funkcijos nustato kiekvienos zonos geometriją (pvz., plotą, perimetrą, storį, centroidą ir pan.). Atliekant zoninę operaciją su dviem rastrais, zoninės funkcijos apdoroja įvesties rastro gardeles zonomis, apibrėžtomis zonų rastro sluoksnyje. Zonų sluoksnyje apibrėžiamos zonos (formos, reikšmės ir vietos), o įvesties rastrė yra reikšmės, naudojamos zonų išvesties reikšmėms apskaičiuoti. 2.33 pav. pavaizduoti dviejų rastrų zoninės statistikos operacijos, skirtos nustatyti vidutinį vandens baseinų nuolydį, rezultatai. Zonas nurodo vandens baseinų rastras, o kitas įvesties duomenų rinkinys yra nuolydžio rastras. Gautas rastras vaizduoja vidutinį vandens baseinų nuolydį.



2.33 pav. Zoninės statistikos pavyzdys

2.2.6 Atstumas

GIS galima matuoti atstumus tarp elementų. Paprasčiausią atstumo statistiką galima gauti žemėlapiu sąsajos lange atstumų užklausos įrankiu. Spustelėjus vienoje vietoje, po to – kitoje, apskaičiuojamas ir ekrane parodomas atstumas tarp šių dviejų vietų. Paprastai šie užklausos įrankiai taip pat gali apskaičiuoti ekrane nubrėžtos polilinijos ilgį.

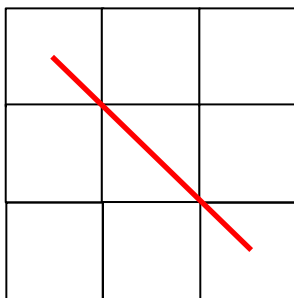
Vektoriniuose duomenų rinkiniuose atstumai paprastai matuojami tarp taškinių elementų; šie elementai gali būti tame pačiame arba skirtinguose duomenų rinkiniuose. Pavyzdžiui, atstumus tarp miestų galima išmatuoti ir įrašyti kaip miestų padėties atributus. Taip pat galima matuoti atstumus nuo taško iki artimiausio kurios nors linijos (polilinijos arba poligono kraštinės) taško. Šio tipo skaičiavimai leidžia nustatyti, pavyzdžiui, mažiausią atstumą nuo pakrantėje esančios vandens mėginių ėmimo vietos iki kranto.

Atstumai jau patys savaime yra naudingas dalykas, tačiau juos galima naudoti ir kaip kintamuosius kuriant modelius. Pavyzdžiui, daugumai gyvūnų rūšių ypač svarbus atstumas iki vandens, todėl kuriant gyvūnų buveinių modelį vietovės tinkamumui įvertinti atstumas iki artimiausio gėlo vandens telkinio (pvz., upės ar ežero) yra svarbus įvesties parametras. Atliekant tokio tipo analizę, rastrinių duomenų modelių grindžiamos funkcijos duoda daugiausia naudos. Pavyzdžiui, kurdami laukinių gyvūnų buveinių modelį, galime sudaryti paviršių, nurodantį atstumą iki gėlo vandens telkinių. Toliau išsamiai aptarsime rastrinės aplinkos atstumo funkcijas.

Atstumas tiesia linija

Atstumas tiesia linija yra fizinis atstumas tarp dviejų taškų. Jis dar vadinamas Euklido atstumu. Rastriniame duomenų rinkinyje atstumai tiesia linija matuojami tarp gardelių centrų (2.34 pav.).

(1, 1)



(3, 3)

2.34 pav. Atstumo tarp rastro gardelių skaičiavimas

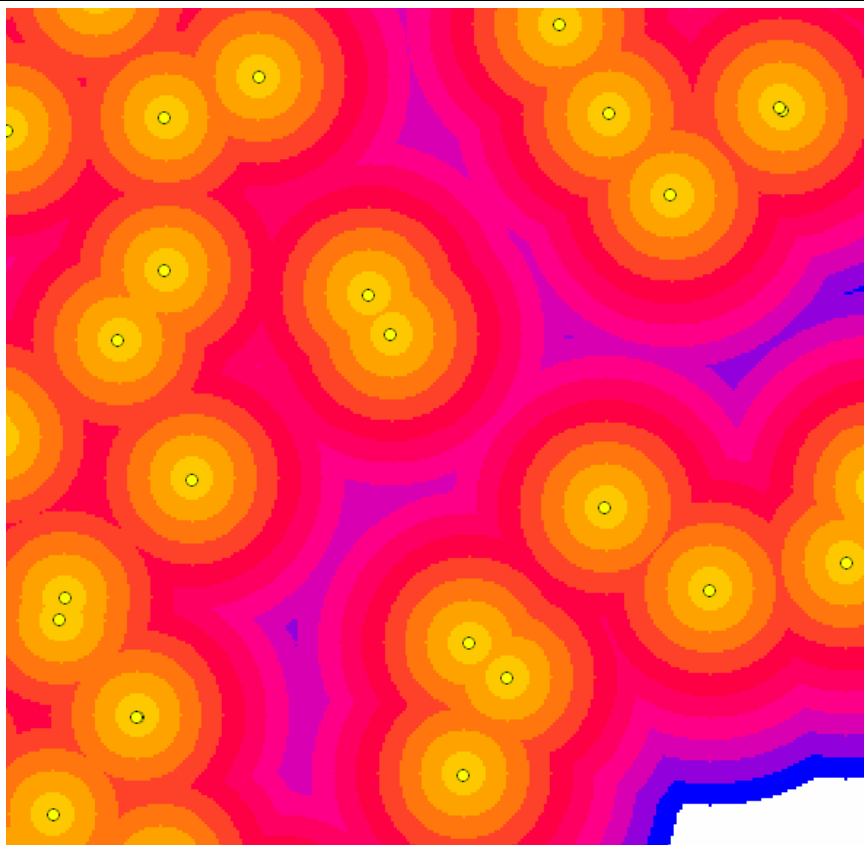
34 pav. atstumas matuojamas tarp gardelių, kurių koordinatės yra 1, 1 ir 3, 3 (sujungtos raudona linija). Jis skaičiuojamas pagal formulę

$$\text{gardelės dydis} \times \sqrt{(3 - 1)^2 + (3 - 1)^2}.$$

Jei gardelės dydis yra 10 metrų, atstumas tarp dviejų taškų bus 28,3 metro, nes:

$$\begin{aligned} &10 \times \sqrt{(3 - 1)^2 + (3 - 1)^2} \\ &10 \times \sqrt{2^2 + 2^2} \\ &10 \times \sqrt{4+4} \\ &10 \times \sqrt{8} \\ &10 \times 2,8284 = 28,2843 \end{aligned}$$

Šiuo pavyzdžiu parodėme, kaip atstumai skaičiuojami, bet atstumo tiesia linija funkcijomis galima gauti ir ištisus išvesties rastrus, kuriuose bus nurodyti atstumai nuo kiekvienos gardelės iki įvesties gardelių arba taškų. Tokia analizė leidžia nustatyti ryšius tarp įvairių vietų. Pavyzdžiui, 2.35 pav. pavaizduotas atstumus tarp miestų (geltoni taškai) vaizduojantis rastras. Šio tipo informacija skirta padėti priimti sprendimus ir padeda planuotojams nustatyti infrastruktūros objektų vietas (pavyzdžiui, jei ligoninė arba poilsio įstaiga galėtų aptarnauti kelias nedideles bendruomenes, jai reikia išrinkti geriausią vietą – tai galima padaryti atstumo funkcijomis).



2.35 pav. Rastrinis atstumo tiesia linija pavyzdys

Paskirstymas ir kryptys

Fizinio atstumo funkcijomis galima generuoti ir paskirstymo bei kryptų rastrus. Paskirstymo rastrų gardelėms priskiriamos artimiausių įvesties gardelių reikšmės. Paskirstymo rastrų pritaikymo pavyzdys gali būti esamų mokyklų mikrorajonų įvertinimas. Sukūrus paskirstymo rastrą, galima nustatyti, kuri mokykla yra arčiausiai kiekvienos žemėlapis vietos. Krypties rastrai vaizduoja kryptį (azimutą) nuo kiekvienos gardelės į artimiausią įvesties objektą. Ši funkcija leidžia paprastai nustatyti kryptį į artimiausią įvesties elementą (pvz., miestą arba infrastruktūros objektą).

Sąnaudų atstumai

Sąnaudų (*cost weighted*) atstumų analizės funkcijos skaičiuoja keliavimo vietove sąnaudas. Šio tipo analizė laidžia nustatyti lengviausią maršrutą tarp dviejų vietų, atsižvelgiant ne tik į fizinį atstumą, bet ir į kitus veiksnus. Pavyzdžiui, trumpiausias kelias tarp dviejų vietovių gali eiti per kalnus; taigi, nors išvengus kalnų fizinis atstumas ir padidėtų, kelionės ilgesniu maršrutu kaina gali būti mažesnė. Skaičiuojant sąnaudų atstumus, dažniausiai atsižvelgiama į daug veiksnių, vaizduojamų atskirais rastrais. Apskaičiuotos sąnaudos yra įvesties rastrinių sluoksnių suma. Vienų rastrų atributų reikšmės nurodo sąnaudų sumažėjimą (vadinamieji teigiami elementai), kitų – padidėjimą. Sąnaudų reikšmės gali būti absoliučios arba santykinės – nurodančios vienų veiksnių įtakos stiprumą lyginant su kitais (didesnės reikšmės atitinka didesnes sąnaudas). Santykinės sąnaudų reikšmės patogios tuo, kad galima suvienodinti įvairių sąnaudų skalę.

Sąnaudų rastrai sudaromi iš duomenų apie sąnaudas, susijusias su analizuojamais veiksniais. Kiekvienam veiksniai sukuriama atskiras rastras, nurodantis konkretaus tipo sąnaudas kiekvienoje gardelėje. Tada vietinės statistikos funkcija galima sugeneruoti

išvestinį (išvesties) rastrą, kurio kiekviena gardelė yra atitinkamų kiekvieno veiksnio rastrų gardelių suma.

Trumpiausias kelias

Turint sąnaudų rastrą, galima nustatyti bendras konkretaus maršruto sąnaudas ir rasti trumpiausią kelią (mažiausių sąnaudų kelią). Šios sąnaudos skaičiuojamos sudedant visas su atstumais tarp gretimų gardelių (horizontaliais, vertikaliais ir įstrižais) susijusias sąnaudas. Kad nustatytume trumpiausią kelią, pirmiausia turime apskaičiuoti sąnaudas tarp visų susijusių gardelių. Tada reikia pasirinkti pradinę gardelę (maršruto pradžios tašką), nuo kurios bus skaičiuojamos sąnaudos judant į gretimas gardeles. Gretima gardelė su mažiausiomis sąnaudomis ir bus kita maršruto gardelė. Tada prie visų sąnaudų bus pridėtos judėjimo į šios gardelės kaimynes sąnaudos. Kartojant šiuos veiksmus, gaunamas trumpiausias kelias.

2.3 Duomenų apibendrinimas

Duomenis galima apibendrinti fiziškai (pvz., pakeisti elementų formas ar išvaizdą), arba galima perklasifikuoti ir apibendrinti jų atributus (pvz., apjungti atributus į bendresnes klases). Tolesniuose skyriuose aptarsime kai kurias vektorinių ir rastrinių duomenų rinkinių apibendrinimo technikas.

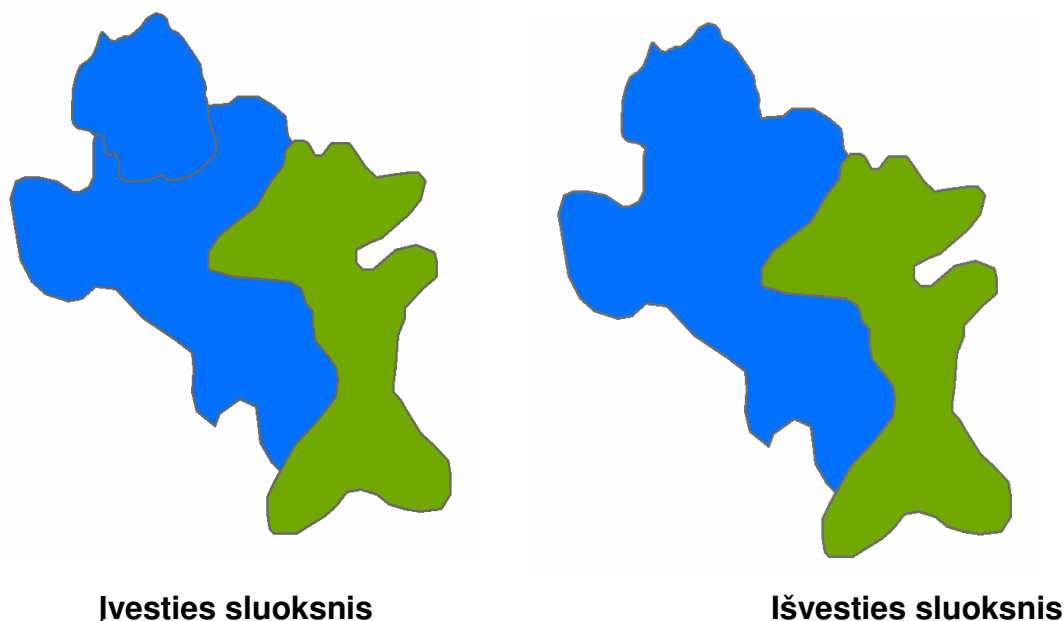
2.3.1 Vektoriniai duomenys

Suliejimas

Suliejimo (*dissolve*) funkcija sujungia elementus, kurių atributai vienodi (2.36 pav.). Šią funkciją galima naudoti sprendžiant dvi svarbias užduotis.

- Ji leidžia supaprastinti klasifikuotą duomenų rinkinį – paversti jį duomenų rinkiniu su bendresnėmis klasėmis. Pavyzdžiui, žemėnaudos poligonus galima apjungti pagal žemėnaudos tipo reikšmes (pvz., kelius, pastatus, grindinius apjungti į vieną universalią klasę – užstatytą teritoriją).
- Ja galima pašalinti ribas tarp vienodą tam tikro atributo reikšmę turinčių poligonų. Pavyzdžiui, sujungus du skirtingos kilmės duomenų rinkinius, gautame rezultate gali būti daug atplaišų arba persidengiančių reikšmių. Suliejimo funkcija gali pašalinti iš šio duomenų rinkinio nereikalingas ribas.

Taip pat šia funkcija galima sumuoti duomenis, nes sujungiant poligonus galima agreguoti ir matematinius duomenis. Pavyzdžiui, suliejant anksčiau minėtas klases į užstatytos teritorijos žemėnaudos klasę, iš gauto sulieto sluoksnio lengva nustatyti visą užstatytos teritorijos elementų plotą. Be to, galima agreguoti ir kitus susijusius atributus (gyventojų skaičių, metines pajamas, gimstamumą ir pan.). Pavyzdžiui, jei miesto seniūnijoms priskirti metinių pajamų atributai, suliedami įvesties elementus galime sumuoti ir vidurkinti jų atributus ir apskaičiuoti viso miesto (arba tam tikros seniūnijų grupės) bendras ir (ar) vidutines metines pajamas.



2.36 pav. Suliejimo pavyzdys

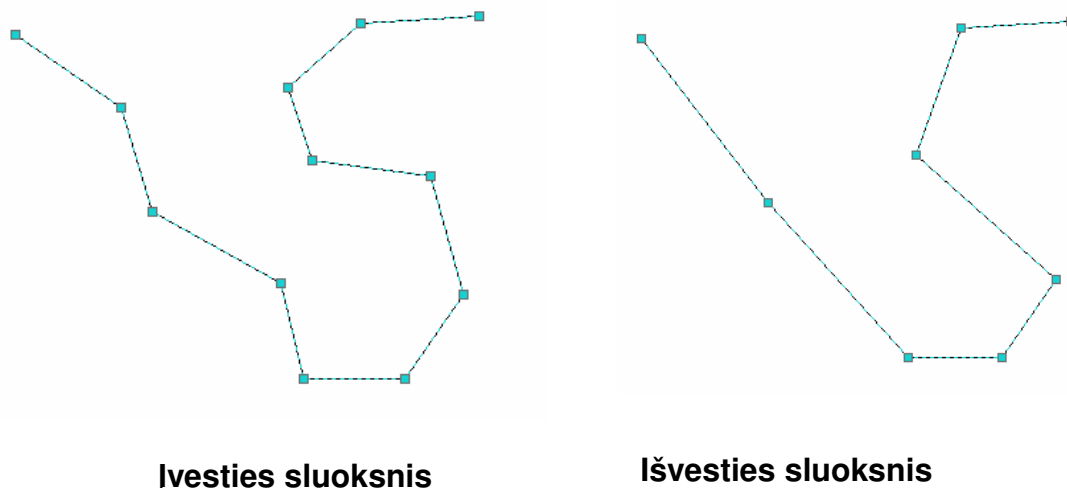
Likvidavimas

Kita duomenų apibendrinimo priemonė yra likvidavimo funkcija – ji sujungia išrinktus poligonus (užklauso rezultata arba išranką) su gretimais poligonais ir sukuria naują sluoksnį. Likvidavimo funkciją išsamiai aptarėme 2.1.2 skyriuje.

Linijos paprastinimas

Iš paprastinamos (*simplify*) linijos pašalinami maži įlinkiai. Tai atliekama pašalinant kai kuriuos linijos taškus (viršūnes). Taip gaunamas apibendrintas linijos elementas (2.37 pav.), išlaikantis bendrą pradinės linijos formą. Linijos paprastinamos tam, kad žemėlapis geriau atrodytų – pavyzdžiui, 1:20 000 masteliu įskaitmenintoje linijoje gali būti tiek taškų, kad atspausdinta 1:100 000 masteliu ji atrodys neaiški. Be to, kuo sudėtingesnė linija, tuo ilgiau trunka jos analizė (pvz., buferių kūrimas), o elementų supaprastinimas prieš buferių analizę dažnai neturi įtakos gauto buferio poligono tikslumui.

Kad supaprastinta linija būtų panaši į pradinę, reikia nurodyti tinkamą paprastinimo laipsnį. Tai atliekama nustatant duomenų mastelį atitinkančią analizės toleranciją. Pavyzdžiui, jei tikslas yra pagerinti duomenų vaizdą žemėlapyje, tolerancijos reikšmė turi būti lygi mažiausiam leistinam tarpui tarp grafinių elementų arba už jį didesnė. Kol rasite visiems rinkinio elementams tinkančią tolerancijos reikšmę, gali tekti atlikti keletą bandymų.



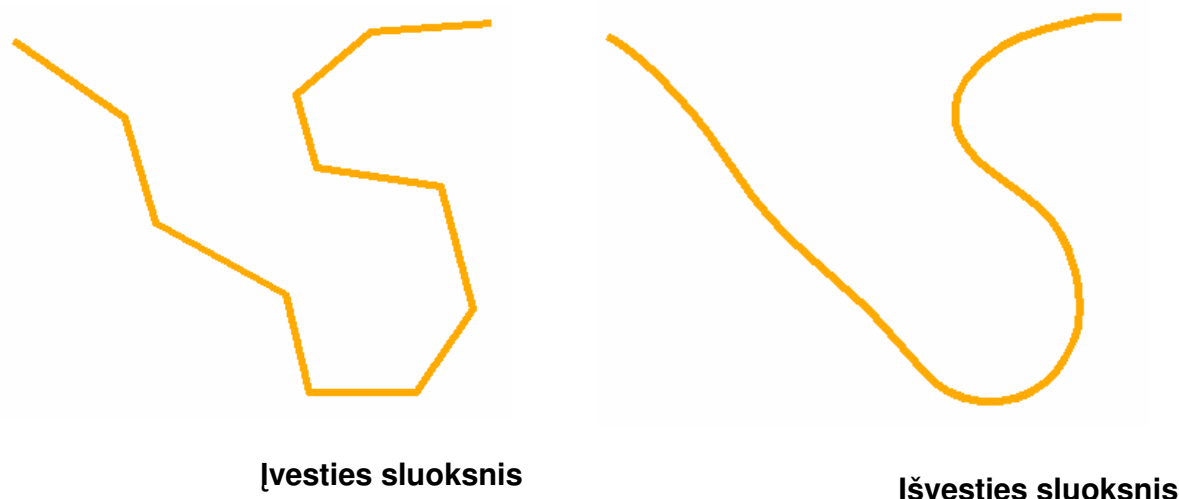
2.37 pav. Linijos paprastinimo pavyzdys

Linijų glodinimas

Linijos glodinamos (*smooth*) tam, kad geriau atrodytų žemėlapyje – pavyzdžiui, pašalinant dantytas dalis pagražinamas linijų estetiškas vaizdas (2.38 pav.). Tai atliekama performuojant liniją matematinėmis formulėmis, sukuriančiomis naujus linijos taškus (viršūnes). Yra įvairiausių glodinimo algoritmų, pavyzdžiui:

- 1 aproksimavimas polinomis eksponentiniu branduoliu (*Polynomial Approximation with Exponential Kernel* - PAEK) – glotni kreivė gaunama tolygaus parametrinio vidurkinimo metodu. Papildomi taškai įterpiami vietose, kurių koordinatės yra svertiniai visų pradinės linijos taškų koordinatės vidurkiai. Artimesniems einamajai koordinatei linijos taškams priskiriami didesni svoriai negu tolimesniems. Gautoje linijoje gali nelikti nė vienos pradinės linijos viršūnės, išskyrus pradinį ir galinį taškus.
- 2 Interpoliavimas Bezje kreivėmis – kiekvienam linijos segmentui pritaikoma Bezje kreivė, ir šios kreivės sujungiamos viršūnėse Beselio liestinės metodu.

Kad gautas rezultatas būtų geras, kaip ir paprastinant linijas, glodinant reikia nurodyti tinkamą analizės toleranciją. Glodinimo atveju tolerancija atitinka ilgį išilgai linijos judančio segmento, pagal kurio taškus skaičiuojamos „glodintos“ koordinatės – kuo didesnė tolerancija, tuo ilgesnis šis segmentas (t. y. tuo daugiau taškų naudojama rezultatui apskaičiuoti) ir tuo glotnesnė linija. Ir šiuo atveju gali tekti atlikti kelis bandymus, kol rasite tolerancijos reikšmę, kuriai esant neprarandamos svarbios linijos charakteristikos (pvz., įlinkiai).



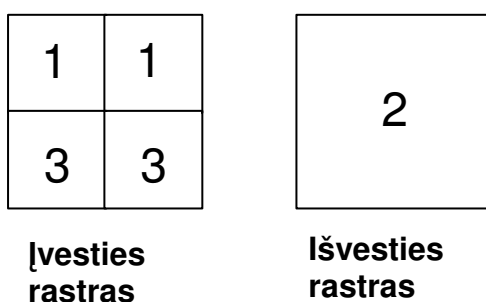
2.38 pav. Linijos glodinimo pavyzdys

2.3.2 Rastriniai duomenys

Kartais rastrinių duomenų rinkinyje gali būti konkrečiai analizei pernelyg išsamių arba nereikalingų duomenų. Pavyzdžiui, iš palydovinio vaizdo gautame žemės dangos duomenų rinkinyje gali būti įvairių mažų gardelių grupių, kurios yra klaidingai klasifikuotos arba per mažos, kad būtų statistiškai reikšmingos. Šių tipų sričių šalinimą iš rastrinio duomenų rinkinio galima automatizuoti apibendrinimo metodais.

Agregavimas

Agregavimas yra performavimo technika, iš įvesties rastro sukurianti mažesnės skiriamosios gebos rastrą (sudarytą iš stambesnių gardelių). Išvesties gardelių reikšmės gali būti skaičiuojamos į jų aprėptį patenkančių įvesties rastro gardelių vidurkio, medianos, sumos, minimumo ar maksimumo pagrindu. 2.39 pav. pateiktas agregavimo analizės pavyzdys, kai išvesties rastro gardelės skaičiuojamos įvesties rastro gardelių vidurkio pagrindu.



2.39 pav. Agregavimo pavyzdys

Agregavimą galima naudoti, pavyzdžiui, tada, kai reikia apibendrinti smulkų temperatūros reikšmių rastrą ir gauti paprastesnį didesnių sričių vidutinės temperatūros rastrą.

Ribų valymas

Ribų valymo (*Boundary clean*) funkcija tinka riboms tarp sričių glodinti. Ji valo ribas gana plačiu mastu, apdorojama duomenis keliais etapais – pirmuoju etapu nagrinėjamos už srities ribų esančios gardelės, antruoju – esančios srities viduje. Praktiškai visų didesnių negu trys kart trys gardelės sričių reikšmės atnaujinamos atsižvelgiant į aplinkinių gardelių reikšmes. „Išoriniu etapu“ už apdorojamos srities esančioms sritims (vienai gardelei kiekviena kryptimi) priskiriamos reikšmės, nustatomos atsižvelgiant į gretimų didesnės svarbos zonų reikšmes. „Vidiniu etapu“ apdorojamos gardelės, kurios nėra visiškai apsuptos tos pačios reikšmės gardelėmis – jų reikšmės gali būti pakeistos aplinkinių gardelių reikšmėmis. Atkreipkite dėmesį, kad gali nebelikti siaurų srities dalių – pavyzdžiui, linijinius objektus vaizduojančių sričių (kad ir upių, kurios yra ilgos, bet tik kelių gardelių pločio).

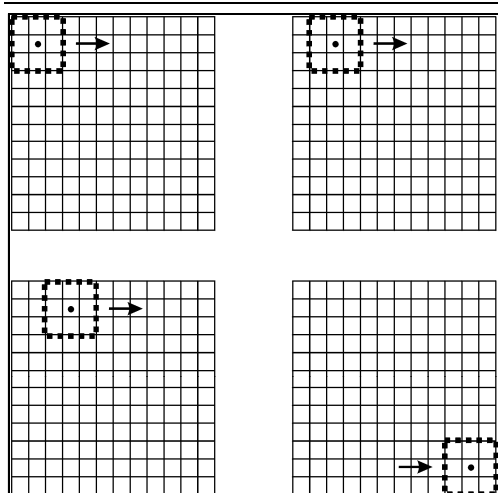
Išplėtimas

Išplėtimo (*expand*) funkcija paplatina nurodytas rastrinio duomenų rinkinio sritis nurodytu gardelių skaičiumi. Žemesnės svarbos gardelės (nurodo naudotojas) pažymimos kaip fono gardelės. Aukštesnės svarbos (priekinio plano) gardeles galima išplėsti į žemesnės svarbos sritis. Ši technika tinka „No Data“ reikšmėms iš duomenų rinkinio pašalinti – vietos be duomenų „užpildomos“ aplinkinių gardelių reikšmėmis.

Filtravimas

Erdvinis filtravimas skirtas konkretiems vaizdo elementams išryškinti arba nuslopinti. Jis pagrįstas erdvinio dažnio charakteristika, susijusia su vaizdo tekstūros sąvoka. Erdvinis dažnis nurodo spalvos kitimo dažnį vaizde. „Stambios“ tekstūros vaizdo dalių, kurių spalva staigiai kinta nedidelėse srityse, erdvinis dažnis aukštas, o „lygių“ dalių, kurių spalva per kelias gardeles pakinta nežymiai, erdvinis dažnis žemas.

Praktiškai filtrai įgyvendinami per įvesties rastrą judančiais langais (branduoliais). Įprastinė filtravimo procedūra yra tokia: kelių gardelių dydžio (pvz., 3×3, 5×5 ir t. t.) „langas“ slenka per visas vaizdo gardeles, ir su į langą patenkančiomis gardelėmis atliekami tam tikri matematiniai skaičiavimai, kurių rezultatas įrašomas į centrinę lango gardelę. Langas stumiamas po vieną gardelę ir eilučių, ir stulpelių ašimis tol, kol perfiltruojamas visas pradinis vaizdas ir sukuriamas naujas. Keičiant skaičiavimo operacijas ir lango gardelių svorius, galima sukurti tam tikro tipo elementus išryškinančius arba nuslopinančius filtrus. Lango judėjimo procesas pavaizduotas 2.40 pav. Paprastai branduolys būna kvadratinis arba stačiakampis, bet gali būti ir žiedinis arba apskritas.



2.40 pav. Lango judėjimo procesas

Filtrai plačiai naudojami įvairių tipų rastrinių duomenų analizei. Bendro pobūdžio taikymo pavyzdžiai yra ribų nustatymas, užlyginimas (glodinimas) ir triukšmų šalinimas. Triukšmai yra klaidingos duomenų reikšmės arba smailės, kurias iš duomenų galima pašalinti. Pavyzdžiui, skaitmeninio aukščių modelio (DEM) smailes galima pašalinti 3×3 medianos filtru.

Teminio filtrų taikymo pavyzdžiai yra paviršiaus šlaito ir ekspozicijos skaičiavimas pagal skaitmeninį reljefo modelį, svertinių funkcijų skaičiavimas atliekant sudėtingą keleto kriterijų rastrinę analizę bei daugelis kitų.

Daugumos (*majority*) filtro funkcija duomenis apibendrina į kiekvieną rastro gardelę įrašydama reikšmes, esančias daugumoje aplinkinių gardelių. Gardelės reikšmė keičiama, jei tenkinamos dvi sąlygos:

- 1 aplinkui turi būti pakankamai daug (pvz., daugiau negu pusė) gardelių, turinčių tą pačią reikšmę;
- 2 tą pačią reikšmę turinčios gardelės turi jungtis erdvėje (sudaryti vientisą sritį) aplink filtro branduolio centrą. Ši sąlyga sumažina tikimybę sugadinti erdvinius duomenų išsidėstymo dėsningumus.

Išrankos taisymas

Išrankos taisymo (*nibble*) funkcija tinka redaguoti rastrinio duomenų rinkinio sritims, kurių reikšmės aiškiai neteisingos arba jų iš viso nėra (pvz., sritims su „No Data“ reikšmėmis). Pirmiausia išrenkamos gardelės, kurių reikšmės bus keičiamos – atliekama užklausa arba sukuriamą išranka. Analizės aprėptis apribojama trafaretu – keičiamos bus į trafaretą patenkančios gardelės. Tada išrinktoms gardelėms paskirstymo pagal Euklido atstumą metodu priskiriamos jų artimiausių kaimynų reikšmės (gardelės paskirstomos pagal mažiausią Euklido atstumą – atstumą tiesia linija).

Regionų grupavimas

Vykdomas skenavimo procesas (panašus į analizę judančiu langu), priskiriantis unikalias reikšmes kiekvienam rastrinio duomenų rinkinio reikšmių regionui (juos sunumeruojant). Gautame duomenų rinkinyje kiekvienam nustatytam regionui priskiriama kita reikšmė (2.41 pav.). Naudotojas negali valdyti priskiriamų reikšmių. Regionų grupavimo funkcija nustato

skirtingus regionus (zonas) ir gali padėti nustatyti objektų išsidėstymo erdvėje dėsningumus.

1	1	1	1
3	3	4	4
3	1	1	1
2	2	2	4

Įvesties rastras

1	1	1	1
2	2	3	3
2	4	4	4
5	5	5	6

Išvesties rastras

2.41 pav. Regionų grupavimo pavyzdys

Sumažinimas

Sumažinimo funkcija (*shrink*) pakeičia klaidingas gardelių reikšmes palei sričių ribas dažniausiomis aplinkinių gardelių reikšmėmis. Sumažinimo funkcija keičiamos gardelės nėra vidinės (t. y. gretimos gardelės jas supa ne iš visų pusių). Atkreipkite dėmesį, kad gali nebelikti siaurų srities dalių – pavyzdžiui, linijinius objektus vaizduojančių sričių (kad ir upių, kurios yra ilgos, bet tik kelių gardelių pločio).

Ploninimas

Ploninimo (*thin*) funkcija leidžia rastriniame duomenų rinkinyje sumažinti linijinius objektus vaizduojančių gardelių skaičių. Pavyzdžiui, nuskenavus popierinį žemėlapi didele skiriamąja geba, linijiniai elementai greičiausiai bus pavaizduoti plačiomis sritimis (pvz., popieriniame žemėlapyje viena linija vaizduojama upė rastre bus 5 gardelių pločio ir 200 gardelių ilgio). Ploninimo funkcija rastro gardeles galima perklasifikuoti taip, kad ši upė taptų 200 gardelių ilgio ir vienos gardelės pločio).

Klausimai žinioms patikrinti

- Kuo skiriasi atributų reikšmėms jungti atliekant perdangą naudojama įnašo taisyklė nuo dominavimo taisyklės?
- Jums reikia nustatyti arčiau negu 500 m ir toliau negu 500 m nuo upių esančių žemdirbystės teritorijų plotų santykį. Jūs turite du vektorinius poligonų sluoksnius: žemdirbystės teritorijų ribų poligonus ir iš upių sluoksniu sukurtus 500 m pločio buferio poligonus. Ar reikiamam santykiui gauti tinka sankirtos (*Intersect*) tipo vektorinė perdanga? Kodėl?
- Kuo skiriasi vietinės, židinio ir zoninės rastrinių sluoksnių operacijos?
- Kam naudotumėte linijinių elementų glodinimo funkciją?
- Ar filtruojant rastrą 3×3 vidurkio (*mean*) filtru kraštutinės reikšmės bus suglodintos, ar išryškintos? Kodėl?

Literatūra

- Chang, K. T. *Introduction to Geographic Information Systems*. McGraw-Hill, 2006.
- Chrisman, N. *Exploring Geographic Information Systems*. John Wiley and Sons, 1997.
- Heywood, I., Cornelius, S. and Carver, S. *An Introduction to Geographical Information Systems*. Pearson Education, 2002.

3. Erdvinė statistika

Kai reikia sumažinti didelį duomenų kiekį iki lengvai suvokiamo informacijos kiekio, dažniausiai pasitelkiama statistinė analizė. Plačiąja prasme apibrėžiama kaip „duomenų surinkimas, analizė, interpretavimas arba aiškinimas bei pateikimas“, statistine analize galima išgauti vertingos informacijos iš milžiniškų duomenų masių.

Nors yra įvairių statistikos formų, mes statistiką skirstysime į klasikinę ir erdvinę. Klasikinė statistika – tai rinkinys metodų, su kuriais susiduria dauguma žmonių – jie naudojami vyriausybės ataskaitose, sporto statistikoje, moksliniuose tyrimuose ir žiniasklaidoje. Kaip ir visi statistiniai metodai, klasikinės statistikos metodai pagrįsti tam tikromis prielaidomis, todėl tinka ne visiems atvejams.

Viena tokių prielaidų yra stebėjimų nepriklausomumas. Ji paprasčiausiai reiškia, kad stebėjimai turi būti nesusiję tarpusavyje: iš vieno stebėjimo gautoje informacijoje nėra jokios informacijos apie kitą stebėjimą. Šis principas dažniausiai pažeidžiamas erdvinio duomenų atveju.

Pirmasis geografinio dėsnio, paprastai vadinamas Toblerio dėsniu, teigia, kad „visa su viskuo susiję, bet artimesni dalykai susiję labiau negu tolimesni“ (Tobleris, 1970: 236). Tai reiškia, kad iš informacijos apie kurią nors vietą galima gauti informaciją apie kitą vietą – ir pastarosios informacijos vertė tuo mažesnė, kuo toliau šios vietos viena nuo kitos.

Siekiant išspręsti šį prieštaravimą, buvo sukurta visa statistikos šaka, atsižvelgianti į stebėjimų priklausomumą. Su kai kuriais erdvinės statistikos metodais susipažinsime šioje kurso dalyje, kaip pavyzdį nagrinėdami savivaldybių lygmens nusikalstamumo Lietuvoje duomenis.

Šią kurso dalį sudaro šešios erdvinės statistikos temos:

- Klasikinė koreliacija, autokoreliacija erdvėje ir vietos koeficientas
- Kintamo plotinio vieneto problema ir ekologinė klaida
- Erdvinio išsidėstymo analizė
- Kraštų įtaka
- Tankio skaičiavimas ir karštųjų zonų nustatymas
- Vietinė erdvinė statistika

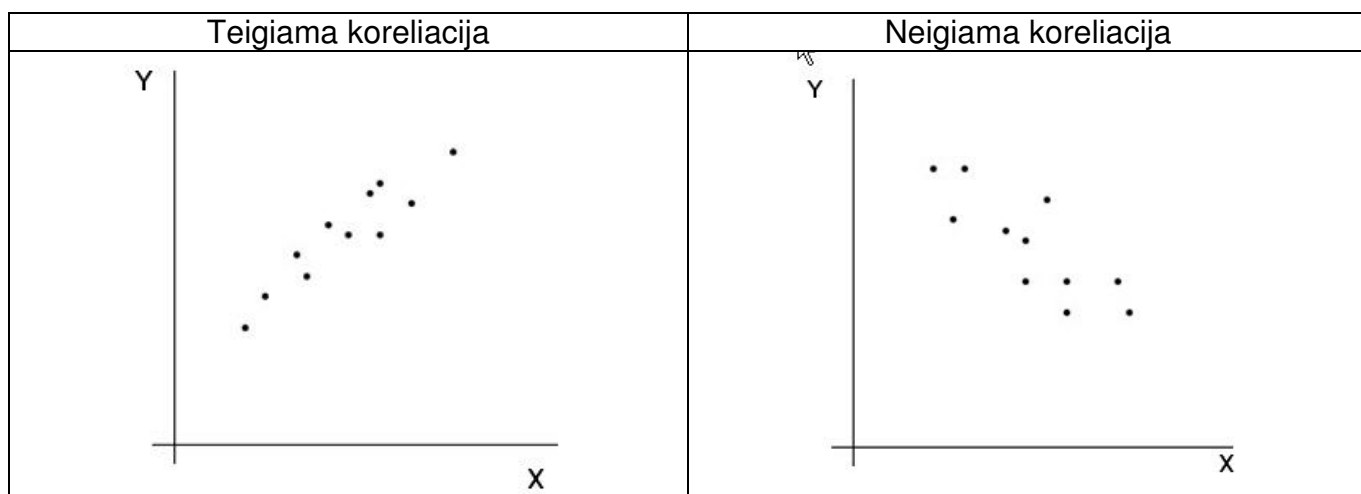
3.1 Klasikinė koreliacija, autokoreliacija erdvėje ir vietos koeficientas

3.1.1 Klasikinė koreliacija

Ankstesnėse kurso dalyse nagrinėjome vidurkio ir standartinio nuokrypio skaičiavimus – t. y. matuojamo dydžio kitimo mastą ir vidutinę reikšmę. O koreliacija yra ryšio tarp dviejų kintamųjų matas. Tirdami koreliaciją, jūs iš esmės klausiate, pavyzdžiui: kaip pasikeis nusikalstamumo lygis, pakitus policijos aktyvumui? Kaip ir turėtų būti, šie du kintamieji gan stipriai susiję – kuo didesnis nusikalstamumas, tuo daugiau reikia policijos pareigūnų; bet kad galėtume tinkamai interpretuoti klasikinės koreliacijos rezultatus, turime išsiaiškinti keletą sąvokų.

Pirma: analizuodami duomenis klasikinės koreliacijos metodais, jūs bandote išsiaiškinti, *AR* du kintamieji susiję. Jei du kintamieji koreliuoja, tai dar nereiškia, kad jie susiję iš tikro. Vienas ekonomistas parašė humoristinį straipsnį, kuriame parodė, kad infliacijos lygis valstybėje stipriai susijęs su dizenterija – taigi koreliaciją reikia interpretuoti atsargiai, ir kintamieji, kurių koreliaciją skaičiuojate, *turėtų* būti tarpusavyje susiję.

Antra: jei du kintamieji kinta ta pačia kryptimi (abu arba didėja, arba mažėja), sakoma, kad jų koreliacija teigiama. Gal kiek neįprasta galvoti, kad du kintamieji koreliuoja teigiamai, kai jie abu mažėja, bet tokia jau šios srities terminologija. Jei vienas kintamasis didėja, kitas mažėja, sakoma, kad kintamieji koreliuoja neigiamai. Du minėti koreliacijos tipai pavaizduoti 3.1 pav.



3.1 pav. Klasikinė koreliacija

Klasikinė koreliacija skaičiuojama taip:

$$r = \frac{\left(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / n \right)}{S_x S_y}$$

r vadinamas koreliacijos koeficientu, x_i ir y_i yra atskirų stebėjimų rezultatai, $\bar{x}$ ir $\bar{y}$ su brūkšneliais viršuje – šių kintamųjų vidurkiai, n – jų imties dydis, o S_x ir S_y – jų standartiniai nuokrypiai. Pagal šią formulę gaunamas skaičius nuo -1 iki +1; teigiamos reikšmės reiškia teigiamą koreliaciją, neigiamos – neigiamą koreliaciją. Jei gaunamas 0, vadinasi, koreliacijos nėra. Jei norite gilintis į šią temą, informacijos galite rasti internete – kelios nuorodos pateiktos papildomos informacijos skyriuje.

3.1.2 Autokoreliacija erdvėje

Autokoreliacija erdvėje (*spatial autocorrelation*) yra erdvinis klasikinės koreliacijos atitikmuo; pagrindinis skirtumas tas, kad autokoreliacijai erdvėje skaičiuoti naudojami stebėjimų rezultatai visada yra aiškiai erdviniai. Užduot klausus, kas atsitinka su vienu kintamuoju kintant kitam kintamajam, skaičiuojant autokoreliaciją erdvėje klausiama, kiek gretimi erdviniai vienetai panašūs vienas į kitą tam tikro kintamojo atžvilgiu. Pavyzdžiui, jei vienoje savivaldybėje gyventojų pajamos didelės, ar gretimose savivaldybėse jos taip pat didelės?

Didžiausias skirtumas tarp klasikinės koreliacijos ir autokoreliacijos erdvėje skaičiavimo tas, kad skaičiuojant pastarąją būtina nurodyti erdvinių vienetų kaimynus. Ar du erdviniai vienetai yra kaimynai, dažniausiai nustatoma vienu iš dviejų metodų: atstumo (*distance*) metodu ir gretimumo (*contiguity*) metodu. Atstumo metodas dažniausiai naudojamas analizuojant taškų išsidėstymą. Jo principas paprastas – jei atstumas tarp taškų nedidesnis už nurodytą, taškai yra kaimynai. Nurodomas atstumas priklauso nuo tiriamo reiškinių ir tyrimų konteksto. Gretimumo metodu tikrinama, ar du plotiniai erdviniai vienetai yra šalia vienas kito. Pavyzdžiui, Lietuva ir Lenkija turi bendrą valstybės sieną, taigi jos gretimos. Šiuo atveju svarbi charakteristika yra gretimumo eilė (*order of contiguity*). Ją paaiškinsime pavyzdžiu (3.2 pav.).

A2	A2	A2	A2	A2
A2	A1	A1	A1	A2
A2	A1	A	A1	A2
A2	A1	A1	A1	A2
A2	A2	A2	A2	A2

2 pav. Erdvinių kaimynų nustatymas

Dominantis erdvinis vienetas pažymėtas raidėmis „A“. Visų kitų erdvinių vienetų žymėjimas atitinka jų gretimumo vienetai „A“ eilę. Visi bendrą ribą su „A“ turintys erdviniai vienetai, net jei ta bendra riba yra tik kampinis taškas, laikomi 1-os gretimumo eilės vienetais ir pažymėti „A1“. Tai vadinama „valdovės gretimumu“ – pagal šachmatų valdovės figūros judėjimą. Erdviniai vienetai, tarp kurių ir vieneto „A“ yra dvi ribos, laikomi 2-os gretimumo eilės vienetais ir pažymėti „A2“. Taip galima plėstis kiek tik reikia, bet daugiausia tyrimams naudojamas 1-os eilės gretimumas – ypač aprašomojoje analizėje, pavyzdžiui, skaičiuojant autokoreliaciją erdvėje.

Tačiau kitas gretimumo matas – „bokšto (*Rook*) gretimumas“ – nelaiko gretimais erdvinių vienetų, kurių bendra riba yra tik kampiniame taške. Valdovės gretimumo principas dažniausiai taikomas vektoriniams duomenims, kadangi dauguma socialinių – ekonominių

– politinių erdvių vienetų nėra kvadratiniai ar stačiakampiai. Be to, valdovės gretimumas nepažeidžia pirmojo (Toblerio) geografijos dėsnio. Taip pat svarbu pažymėti, kad yra ir kitų gretimumo matų: kaimynai gali būti nustatomi pagal erdvių vienetų bendrų ribų ilgį arba pagal dvikrypčius svėrų koeficientus, nurodančius individų srautą iš gretimų erdvių vienetų ir į juos; nuo pasirinkto mato priklauso erdvinės analizės rezultatai. Pavyzdžiui – vėl apie valstybes, – bendra Lietuvos ir Latvijos valstybės siena ilgesnė negu Lietuvos ir Lenkijos, bet Lenkijos ekonomikos apimtis daug didesnė. Taigi, kaimynystės matas (ir kaimynų svarba) priklauso nuo konteksto. Tačiau dažniausiai analizei naudojamas pirmosios eilės valdovės gretimumas.

Dažniausias autokoreliacijos erdvėje matas yra Morano I :

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^n (y_i - \bar{y})^2 \right) \left(\sum_{i \neq j} \sum w_{ij} \right)},$$

kur kintamieji n ir y yra tie patys kaip ir ankstesnėje formulėje, o w_{ij} yra erdvių vienetų erdvių kaimynų matrica. Ši formulė gana sudėtinga ir ją suprasti kol kas nėra būtina, nes dauguma programų, tarp jų ir ArcGIS, atliks šiuos skaičiavimus už jus. Svarbu atkreipti dėmesį, kad Morano I ir klasikinės autokoreliacijos formulės panašios. Pagrindinis skirtumas tas, kad skaičiuojant autokoreliaciją erdvėje analizuojamas tik vienas kintamasis. Autokoreliacija erdvėje – tai ne kintamųjų, o erdvių vienetų koreliacija.

Morano I reikšmių intervalas yra nuo -1 iki $+1$, kaip ir klasikinės koreliacijos. Tai ir yra pagrindinė priežastis, dėl kurios Morano I populiarnesnė už Gyrio (*Geary*) C – pastarasis autokoreliacijos erdvėje matas skaitinėmis reikšmėmis nėra toks panašus į klasikinę koreliaciją. Jei Morano I lygi nuliui, reiškia, kad autokoreliacijos erdvėje nėra, jei didesnė už nulį – autokoreliacija erdvėje teigiama, jei mažesnė už nulį – autokoreliacija erdvėje neigiama. Taigi, jei autokoreliacija erdvėje teigiama (dažniausias atvejis tiriant socialinius – ekonominius – politinius erdvių vienetų), tai kaimyninių erdvių vienetų reikšmės panašios. O jei autokoreliacija erdvėje neigiama, kaimyninių erdvių vienetų reikšmės nepanašios.

3.1.3 Vietos koeficientas

Vietos koeficientas (*location quotient*) yra aprašomasis matas, statistikoje pradėtas naudoti gana neseniai. Nuo penktojo dešimtmečio pradėtas naudoti ekonominėje geografijoje užimtumui arba pramonės specializacijai matuoti (Isard ir kiti, 1998; Miller ir kiti, 1991), vietos koeficientas pastaruoju metu naudojamas erdvinėje nusikalstamumo analizėje (Brantingham ir Brantingham, 1995, 1998). Šį koeficientą gerai žino geografi: jis nurodo tam tikro aktyvumo dalies viename erdviniam vienetui santykį su tuo paties aktyvumo dalimi visame regione. Jis atspindi šio aktyvumo padidėjimą arba sumažėjimą ir labai praverčia atliekant bet kokią analizę, skirtą nustatyti regionams, kuriuose tam tikras aktyvumas pasireiškia labiau negu kituose. Vietos koeficientą paaiškinsime pavyzdžiu. Tarkime, kad turime nusikalstamumo Lietuvos savivaldybėse duomenis.

Vietos koeficientas skaičiuojamas taip:

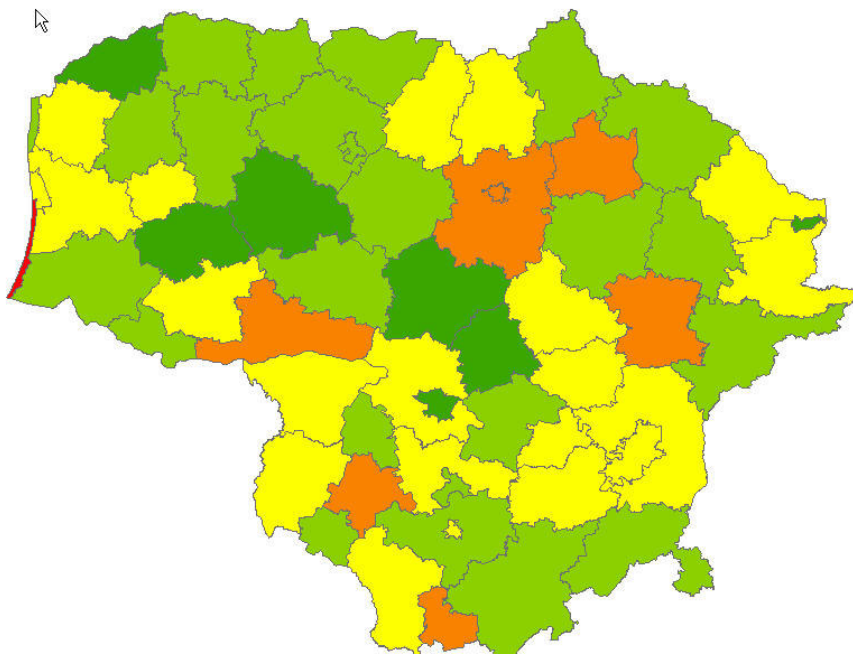
$$LQ = \frac{C_{in}/C_m}{\sum_{n=1}^N C_{in} / \sum_{n=1}^N C_m},$$

kur C_{in} yra i tipo nusikaltimų skaičius erdviniam vienetui n , C_m – visų nusikaltimų skaičius erdviniam vienetui n , o N – erdvių vienetų skaičius. Šiame pavyzdyje vietos koeficientas yra konkretaus nusikaltimų tipo tam tikroje Lietuvos savivaldybėje dalies santykis su to paties nusikaltimų tipo dalimi visoje Lietuvoje. Jei vietos koeficientas lygus vienetui, šiai savivaldybei tenka proporcinga tam tikro tipo nusikaltimų dalis; jei vietos koeficientas didesnis už vienetą, šio tipo nusikaltimų savivaldybėje neproporcingai daug; jei vietos koeficientas mažesnis už vienetą, savivaldybėje šių nusikaltimų dalis neproporcingai mažesnė.

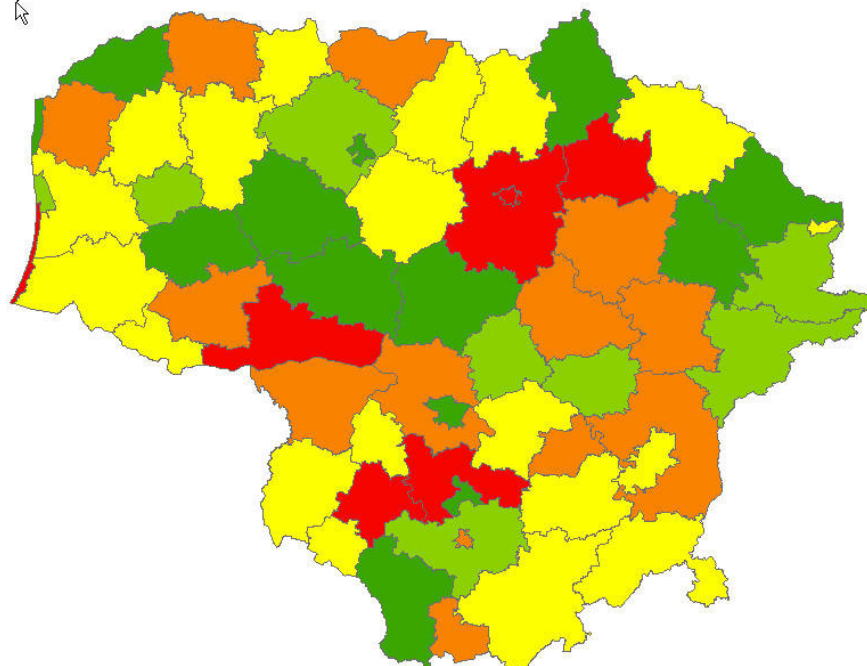
Panagrinėkime dar vieną pavyzdį apie Lietuvos savivaldybes. Kaip pavaizduota 3.3 pav., daugumoje Lietuvos savivaldybių įsilaužimai nėra didelė problema (legenda neįtraukta trumpumo dėlei; žalia spalva reiškia mažą nusikalstamumą, geltona – vidutinį, raudona – didelį). Yra kelios savivaldybės, kuriose įsilaužimų skaičius gana didelis, bet daugumoje Lietuvos savivaldybių jie nėra didelė problema.

Vietos koeficientas parodo kitokį įsilaužimų vaizdą. Aiškiai išsiskiria kelios savivaldybės (raudonos), kuriose šių nusikaltimų dalis didesnė negu visoje Lietuvoje. Nors daugumoje iš jų didesnis ir įsilaužimų skaičius, bet ne visose. Šis nusikalstamumas nestipriai dominuoja (oranžinė spalva) daugelyje savivaldybių, kuriose nusikalstamumas mažas. Tai nereiškia, kad šiose savivaldybėse nesaugu – nusikaltimų daroma visur ir visada, – tai rodo, kad vienos savivaldybės įsilaužimai populiarnesni negu kitose. Ši informacija svarbi formuojant politiką; ji rodo, kad šiose savivaldybėse esamos policijos ir nusikaltimų prevencijos pajėgos daugiau dėmesio turėtų skirti įsilaužimams – tai nereiškia, kad pačių šių pajėgų reikia daugiau. Vietos koeficientą galima naudoti tiriant bet kokią reiškinį, kuris įvairiose vietose vyksta skirtingai – tai yra, praktiškai bet kokią žmonių veiklą.

a) Nusikaltimų skaičius 10 000 gyventojų



b) Vietos koeficientas



3.3 pav. Įsilaužimai Lietuvos savivaldybėse

3.2 Kintamo plotinio vieneto problema ir ekologinė klaida

3.2.1 Kintamo plotinio vieneto problema

Kintamo plotinio vieneto problema (KPVP, *modifiable areal unit problem*) būdinga visiems erdviniams duomenims, kurie agreguojami naudojant poligonus. Pirmasis ją nurodė Oupenšo (*Openshaw*) 1984 metais. Kintamo plotinio vieneto problema reiškia paprastą dalyką – analizės rezultatai priklauso nuo to, kaip pasirenkami erdviniai vienetai. Pavyzdžiui, jei atlikdami kokią nors analizę vieną kartą Lietuvą padalinsime į savivaldybes, kitą kartą – į apskritis, analizės rezultatai būtinai skirsis, o gal net bus visiškai priešingi – taigi KPVP tikrai yra problema. Yra dvi KPVP apraiškos: mastelio efektas ir agregavimo arba zonų formavimo efektas.

Mastelio efektą ką tik aptarėme pavyzdyje apie tą pačią analizę, atliekamą suskirsčius teritoriją į savivaldybes ir apskritis. Paprastai tariant, (erdviniai) statistiniai rezultatai gali priklausyti nuo duomenų erdvinės skiriamosios gebos. Apskritai, kuo didesnis poligonų plotas, tuo labiau tikėtina, kad jiems apskaičiuotos reikšmės bus panašios, taigi, stipriai koreliuos. Taip yra todėl, kad didelis erdvinis vienetas yra jo sudedamųjų dalių vidurkis. Tarkime, kad nedarbo lygis Lietuvos savivaldybėse labai skiriasi, ir kad kiekvienoje apskrityje yra ir aukšto nedarbo lygio, ir žemo nedarbo lygio savivaldybių. Agregavus savivaldybes į apskritis, gali būti, kad visose apskrityse nedarbo lygis bus panašus – vien todėl, kad jo reikšmės tapo suvidurkintos ir skirtumų nebeliko.

Agregavimo arba zonų formavimo efektas pasireiškia tuo, kad statistiniai rezultatai priklauso nuo plotinių erdvių vienetų formavimo metodikos. Dar kartą panagrinėkime Lietuvos apskričių ir savivaldybių atvejį. Kiekviena iš dešimties dabartinių apskričių užima tam tikrą geografinę sritį ir turi tam tikrą plotą. Dabar tarkime, kad „sumaišėme“ apskričių savivaldybes taip, kad kiekvienos apskrities savivaldybių rinkinys pasikeitė, bet apskrities plotas išliko beveik tas pats – gauti lygiai tą patį plotą gali ir nepavykti. Išlaikydami tą patį geografinį apskričių dydį, išvengiame mastelio efekto. Bet apskričių savivaldybės dabar jau kitos, taigi apskričių statistiniai rezultatai taip pat pasikeis.

Trumpai tariant, pakeitę erdvių vienetų skaičių ir struktūrą, galime gauti visiškai skirtingus analizės rezultatus.

3.2.2 Kintamo plotinio vieneto problemos minimizavimas

KPVP atsiranda dėl paprasčiausios priežasties: analizei naudojamus erdvinis vienetus gauname jau apibrėžtus, o juos apibrėžiant nebuvo atsižvelgta į statistinę analizę. Pavyzdžiui, jei Kanadoje reikia pertvarkyti balsavimo apylinkių ribas, jos dažniausiai pertvarkomos taip, kaip patogiau valdančiajai partijai. Tai gali padėti pasiekti tam tikrų politinių tikslų, bet sukuriama „klaidingi“ erdviniai vienetai, pagal kuriuos vėliau atliekama analizė.

KPVP iš tikro minimizuoti galima tik parinkus konkrečiai analizei tinkamus erdvinis vienetus. Kitaip tariant, kai tik žinosime „teisingus“ erdvinis vienetus, KPVP išnyks. Aišku, kad tokio lygio žinios, jei iš viso gali būti, kiekvienai analizei bus skirtingos: nedarbui tirti tinkami plotiniai vienetai būtinai skirsis nuo infarkto rizikai tirti tinkamų plotinių vienetų. Sunkiausia tada, kai žinome, kad ši problema egzistuoja, bet sprendimo dar neradome. Ką daryti tokiu atveju?

Paprasčiausias atsakymas – atlikti kuo daugiau tyrimų, pasirenkant įvairius erdvinis vienetus. Jei visų tyrimų kiekybiniai ir kokybiniai rezultatai tie patys (arba labai panašūs) – pavyzdžiui, visos koreliacijos teigiamos ir maždaug to paties dydžio – tada nėra ko jaudintis. Jei rezultatai kokybiškai panašūs (visais atvejais koreliacijos teigiamos), bet šių rezultatų skaitinės reikšmės skiriasi, tada geriau būti per daug atsargiems, negu daryti nepagrįstas išvadas apie dviejų kintamųjų ryšį. Blogiausiu atveju, jei rezultatai visiškai prieštaringi (koreliacijos tai teigiamos, tai neigiamos, jų reikšmės irgi kinta), tada šiais rezultatais tiesiog negalima pasitikėti. Reikia atlikti papildomą analizę, parinkus kitus erdvinis vienetus, kintamuosius arba ilgesnes laikines sekas. Tuo labiau negalima tokiais rezultatais grįsti politinių sprendimų.

Deja, daugeliu analizės atvejų erdvinis vienetų patiems pasirinkti negalima. Jei taip atsitiktų – o taip būtinai atsitiks, rezultatus reikia interpretuoti atsargiai, atsižvelgti į tai, kad jie gali priklausyti nuo erdvinis vienetų pasirinkimo.

3.2.3 Kintamo plotinio vieneto problema ir ekologinė klaida

Vadinamoji ekologinė klaida (*ecological fallacy*), kaip ir šiame kurse, paprastai nagrinėjama kartu su kintamo plotinio vieneto problema; bet tai kita problema, dažnai kylanti dirbant su geografiniais duomenimis. Pirmasis ją nurodė Robinsonas (*Robinson*) 1950 m. Ekologinė klaida daroma tada, kai iš viena erdvine skiriamąją geba atliktos analizės rezultatų daromos išvados apie kitos erdvinės skiriamosios gebos rezultatus. Pavyzdžiui, įsivaizduokime iš dviejų savivaldybių sudarytą apskritį. Vidutinės gyventojų pajamos per mėnesį apskrityje yra 1 000 litų. Iš šios statistikos galima daryti išvadą, kad ir kiekvienoje savivaldybėje vidutinės gyventojų pajamos per mėnesį yra 1 000 litų. Tačiau iš tikro šios pajamos gali būti 500 ir 1 500 litų.

Šiuo atveju ekologinė klaida redukuojama iki dalinimo klaidos: kas teisinga apie visumą, teisinga ir apie jos dalis. Tačiau pakanka apsidairyti aplink, ir matyti, kad žmonių vidutinės mėnesio pajamos stipriai skiriasi, ir joks vidurkis nieko nesako apie konkretų žmogų. Taigi, kas teisinga apie visumą, nebūtinai teisinga apie jos dalis.

Kaip išvengti ekologinės klaidos? Priešingai negu KPVP, ekologinės klaidos išvengti paprasta – reikia gerai apgalvoti daromas išvadas. Jei analizuojate tik apskritis, išvadas galima daryti tik apie apskritis; tai galioja ir savivaldybių, mikrorajonų bei individų atveju. Moksliniuose tyrimuose, vyriausybiniuose tyrimuose ir žiniasklaidoje ekologinių klaidų daug daugiau, negu norėtumėte.

Ekologines klaidas sunku pastebėti dar ir todėl, kad ne visos jos daromos su erdviniais vienetais – tai tik aiškiausia ir lengviausia aptinkama ekologinės klaidos forma. Žmonių grupėse pagal tam tikras socialines charakteristikas gali išryškėti tam tikra vidutinė tendencija, pavyzdžiui, aukštas intelektas. Tačiau tai nereiškia, kad kiekvienas šios grupės narys labai protingas.

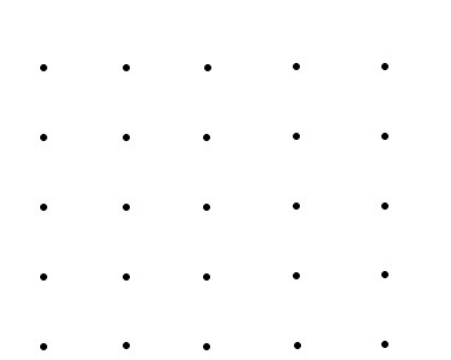
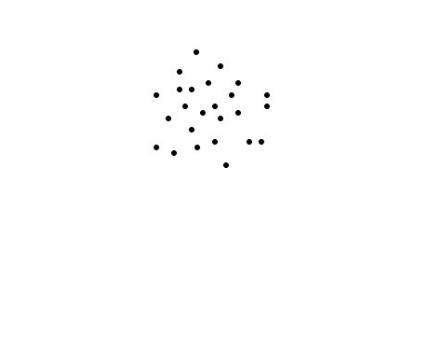
3.3 Erdvinio išsidėstymo analizė

3.3.1 Tolygus, atsitiktinis ir grupinis išsidėstymas

Socialinių ir gamtos reiškinių išsidėstymas erdvėje labai domina ir įvairiausio pobūdžio įstaigų (vyriausybės, mokslo įstaigų, privačių įmonių) darbuotojus, ir įvairiausių sričių specialistus; nors pradžioje laikyta geografijos dalimi, išsidėstymo analizė (*pattern analysis*) dabar naudojama ir ekologijoje (plačiąją prasme – biologijoje), medicinoje, visuomenės sveikatos srityje bei statistikoje.

Tiriant išsidėstymą, išskiriami trys bendrieji išsidėstymo erdvėje tipai: tolygus (*uniform*), atsitiktinis (*random*) ir grupinis (*clustered*) (3.4 pav.). Atkreipkite dėmesį: nors 3.4 pav. pavaizduotas taškų išsidėstymas, ir kalbėdami apie išsidėstymo analizę taip pat manipuliuosime taškais, bet yra išsidėstymo analizės metodų, skirtų poligonais išreikštų erdvių vienetų išsidėstymui analizuoti. Tačiau tradiciškai išsidėstymo analizė yra taškinė, taigi šiame skyriuje nagrinėjama taškų išsidėstymo analizė. Plotinių erdvių vienetų klasterius analizuosime 3.5 temoje.

3.4a pav. pavaizduotas tolygus taškų išsidėstymas. Tolygus taškų išsidėstymas (kartais vadinamas „išsklaidytu“) rodo, kad kiekvieno taško vietą lemia sistemingas erdvinis procesas. Kitaip tariant, žinant vieno taško vietą ir proceso pobūdį, galima tiksliai numatyti kitų taškų vietas. 3.4a pav. pavaizduotu atveju bet kuris taškas bus vienu vienetu aukšty (žemyn) arba kairėn (dešinėn) nuo kurio nors kito taško. Tokio tipo taškinių erdvių išsidėstymų gamtoje aptinkama retai, dažniausiai jie atsiranda dėl inžinerinės žmogaus veiklos. Geriausias šio išsidėstymo pavyzdys yra šviesoforai stačiakampio gatvių tinklo, dažniausiai vadinamo Manhatano tinkleliu, sankirtose taškuose. Tokio tinklo gatvės abiem kryptimis (šiaurės – pietų ir rytų – vakarų) išdėstytos vienodais atstumais, taigi jų sankirtose įrengti šviesoforai erdvėje bus pasiskirstę tolygiai.

a) tolygus	b) atsitiktinis	c) grupinis
		

3.4 pav. Tolygus, atsitiktinis ir grupinis taškų išsidėstymas

Atsitiktiniame taškų erdviame išsidėstyme (3.4b) nematyti jokios vyraujančios tendencijos. Esant tokiam išsidėstymui, vieno taško vieta nieko nesako apie kitų taškų vietas – nustatyti kito taško padėtį kokiais nors matematiniais skaičiavimais tas pat, kaip pataikyti į jį pirštu užsimerkus. Tokius erdvių taškų išsidėstymus sukuria vadinamieji Puasono procesai. Todėl hipotezę apie išsidėstymo atsitiktinumą lengva patikrinti.

3.4c pav. pavaizduotas grupinis erdvinis taškų išsidėstymas. Esant šiam erdviniam išsidėstymui, taškų tankis labai netolygus, nes vienos vietose yra dauguma taškų, o kitose jų beveik nėra. Taigi numatyti kitų taškų padėtis galima, nes tikėtina, kad arti esamo taško bus ir kitų taškų, o arti vietos, kurioje taškų nėra, jų ir nebus. Tokios taškų sankaupos gali vaizduoti, pavyzdžiui, mažmeninės prekybos taškus centriniame verslo rajone arba ligos protrūkį dėl fiksuoto toksinų šaltinio.

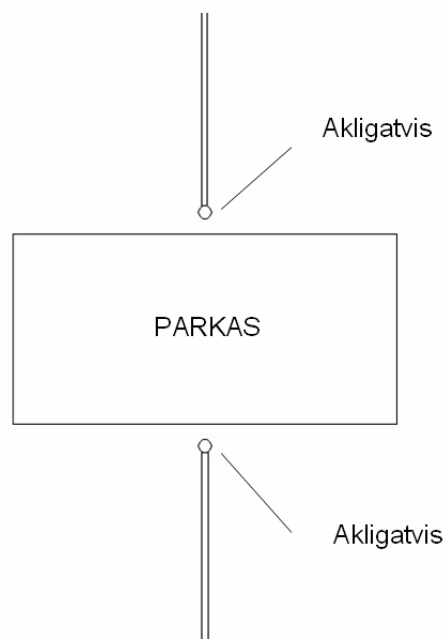
Panagrinėsime dvi išsidėstymo analizės formas. Pirmoji ir dažniausiai naudojama – artimiausių kaimynų analizė: ar taškų erdvinis išsidėstymas atitinka konkretų skirstinį, nustatoma pagal atstumus tarp taškų. Antroji – analizė tinkleliu (*quadrat analysis*): sritis sudalinama į plotinius vienetus, ir išsidėstymo erdvėje forma nustatoma pagal taškų skaičių dalyse.

Atkreipkite dėmesį, kad šios temos literatūroje „tašku“ dažniausiai vadinama bet kuri vieta žemėlapyje, o „įvykis“ yra tašku pažymėta vieta žemėlapyje, kurioje kažkas įvyko: nustatytas ligos atvejis, įvykdytas nusikaltimas, įvyko avarija ir pan. Nors toliau šioje temoje vartosime terminą „taškas“, minėtas skirtumas taps svarbus kitoje temoje, kai nagrinėsime kraštų įtaką.

3.3.2 Artimiausių kaimynų analizė

Atliekant artimiausių kaimynų analizę (*nearest neighbor analysis*), matuojami atstumai nuo taškų iki jų „artimiausių kaimynų“, skaičiuojamas visų duomenų rinkinio taškų atstumų iki artimiausių kaimynų vidurkis, kuris lyginamas su teoriniu (pavyzdžiui, atsitiktinio (Puasono) tikimybės skirstinio) vidutiniu atstumu. Atstumai nuo taškų iki jų artimiausių kaimynų yra Euklido atstumai, t. y. atstumai tiesia linija. Toks atstumo matavimo būdas turi ir pranašumų, ir trūkumų. Pranašumas tas, kad atstumus paprasta apskaičiuoti, naudojantis lygių intervalų tinkleliu ir keliomis geometrijos funkcijomis – tai matysime iš toliau pateikto pavyzdžio. Trūkumas tas, kad Euklido atstumas ne visada tinka analizuojant žmonių gyvenviečių išsidėstymą.

Pavyzdžiui, 3.5 pav. pavaizduoti du akligatviai, tarp kurių yra parkas. Kad būtų paprasčiau, įsivaizduokite, jog iš namo vienoje parko pusėje kitoje parko esančio namo net nematyti, nes jį užstoja medžiai ir pan., bet šie namai yra gan arti vienas kito. Jei bus matuojami Euklido atstumai, gali būti nustatyta, kad šie du skirtingose parko pusėse esantys namai yra artimiausi kaimynai, nors tinklo atstumas (praktiškesnis atstumo matas šiuo atveju) tarp jų gali būti gana didelis. Aišku, artimiausių kaimynų analizei galima naudoti ir tinklo atstumus. Analizės esmė išliks ta pati, tik atstumams skaičiuoti reikės daugiau kompiuterio išteklių. Pateikiamame pavyzdyje naudojami Euklido atstumai.



3.5 pav. Problema su Euklido atstumais

Norint nustatyti erdvinio taškų išsidėstymo tipą, reikia atlikti keletą paprastų skaičiavimų. Pirmiausia apskaičiuojamas vidutinis atstumas iki artimiausio kaimyno:

$$NND^* = \frac{\sum_{i=1}^n NND_i}{n},$$

kur NND_i – atstumas iki artimiausio kaimyno nuo taško i , n – taškų skaičius. Jei taškų išsidėstymas idealiai grupinis (visi taškai vienoje vietoje), NND^* reikšmė yra 0. Tada apskaičiuojamas atsitiktinio erdvinio skirstinio teorinis vidutinis atstumas iki artimiausio kaimyno:

$$NND^R = \frac{1}{2\sqrt{Tankis}},$$

kur $tankis$ lygus taškų skaičiaus ir nagrinėjamos srities ploto santykiui. Tada iš šių dviejų kintamųjų gaunamas normuotas artimiausio kaimyno rodiklis (R):

$$R = \frac{NND^*}{NND^R}$$

pagal kurį galima nustatyti erdvinio taškų išsidėstymo pobūdį. Normuotas artimiausio kaimyno rodiklis R patogus tuo, kad jo reikšmių intervalas griežtai apibrėžtas (0 – 2.149), todėl galima sudaryti tokią klasifikaciją:

3.1 lentelė. Klasifikavimo pagal normuotą artimiausio kaimyno rodiklį R schema

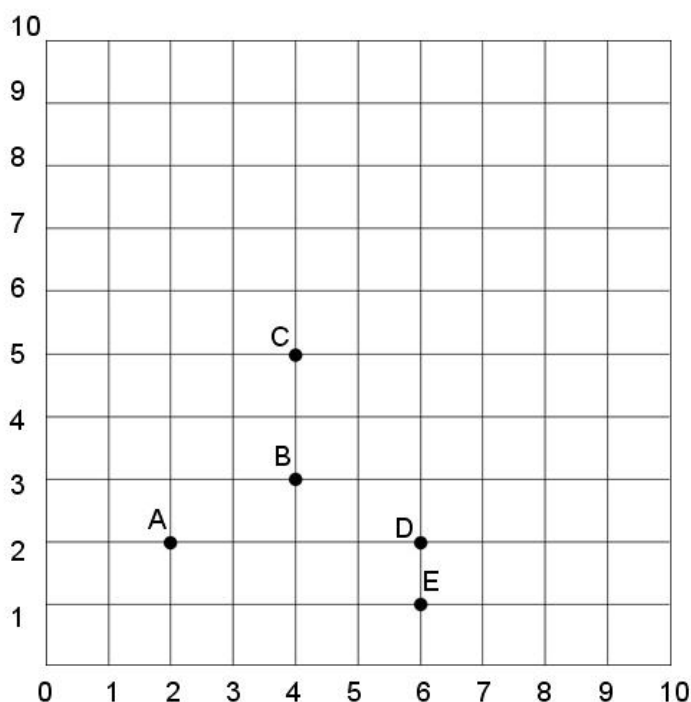
Idealiai grupinis	Labiau grupinis, negu atsitiktinis	Atsitiktinis	Labiau tolygus, negu atsitiktinis	Idealiai tolygus
$R = 0$	$R = 0,5$	$R = 1$	$R = 1,5$	$R = 2,149$

Tokios aprašomosios klasifikacijos analizei dažnai nepakanka, tad norint nustatyti erdvinio taškų išsidėstymo pobūdį reikia atlikti statistinių hipotezių tikrinimą. Tokio tikrinimo nulinė hipotezė yra ta, kad erdvinis taškų išsidėstymas yra atsitiktinis, o tikrinimo statistika yra

$$Z_{NND} = \frac{NND^* - NND^R}{\sigma_{NND}}, \text{ kur } \sigma_{NND} = \frac{0.26136}{\sqrt{n(Tankis)}}.$$

Atliekant vienusi statistinį bandymą su 10 proc. tikimybe (p -reikšmė = 0,10), gaunama kritinė Z_{NND} reikšmė apytiksliai lygi 1,28. Todėl, jei $Z_{NND} > 1,28$, erdvinis taškų išsidėstymas labiau tolygus negu atsitiktinis, jei $Z_{NND} < -1,28$, erdvinis taškų išsidėstymas daug labiau grupinis negu atsitiktinis, o jei $-1,28 < Z_{NND} < 1,28$ – erdvinis taškų išsidėstymas nedaug skiriasi nuo gaunamo atsitiktiniu Puasono erdviu procesu. Artimiausio kaimyno analizės ir susijusios statistinės analizės taikymą iliustruosime pavyzdžiu.

3.6 pav. pavaizduota tiriama sritis su 5 taškais, kurių erdvinis išsidėstymas panašus į grupinį. Tiriamos srities plotas yra $10 * 10 = 100$ vienetų.



3.6 pav. Artimiausių kaimynų analizei naudojamų taškų vietos

Taškai, jų erdvinės koordinatės, artimiausi kaimynai ir atstumai iki artimiausių kaimynų surašyti 3.2 lentelėje.

3.2 lentelė. Koordinatės ir artimiausi kaimynai (NN)

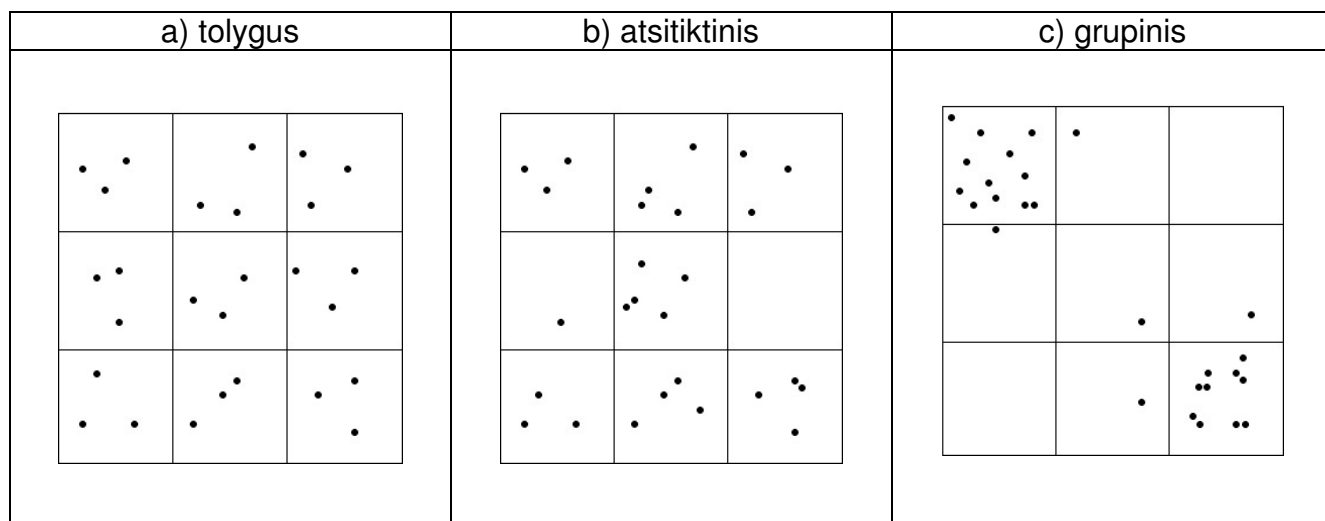
Taškas	X	Y	NN (artimiausias kaimynas)	NND (atstumas iki artimiausio kaimyno)
A	2	2	B	2,24
B	4	3	C	2
C	4	5	B	2
D	6	2	E	1
E	6	1	D	1

Iš šių skaičių lengva gauti likusius kintamuosius: $NND^* = (8,24)/5 = 1,648$, $NND^R = 1 / (2 \cdot \sqrt{5/100}) = 1/(2 \cdot \sqrt{0,05}) = 2,236$, o $R = 1,648 / 2,236 = 0,737$. Kaip ir buvo pastebėta, šis erdvinis išsidėstymas labiau grupinis negu atsitiktinis. Liko nustatyti, ar šios erdvinės grupės reikšmingos statistiškai. Į šį klausimą galima atsakyti apskaičiavus $Z_{NND} = (1,648 - 2,236) / 0,523 = -1,13$. Taigi, pagal šį statistinį bandymą nulinės hipotezės apie atsitiktinį erdvinį išsidėstymą atmesti negalima. Ir nors pavaizduotas išsidėstymas atrodo labiau grupinis negu atsitiktinis, statistiniu požiūriu taip nėra.

3.3.3 Analizė tinkleliu

Kitas metodas erdvinio taškų išsidėstymo pobūdžiui nustatyti yra analizė tinkleliu (*quadrant analysis*). Šiuo atveju skaičiuojami ne atstumai tarp taškų, o taškų dažnis (skaičius) tam tikrame plote. Ant nagrinėjamos srities užklojamas tinklelis (dažniausiai iš kvadratinių gardelių), ir suskaičiuojama, kiek taškų yra kiekvienoje gardelėje. Taškų išsidėstymo erdvėje pobūdis nustatomas analizuojant taškų dažnį gardelėse.

Apskritai kalbant, jei visose kvadratinėse gardelėse taškų skaičius vienodas (kintamumo nėra), laikoma, kad taškai erdvėje išsidėstę tolygiai; jei taškų skaičius gardelėse stipriai skiriasi (didelis kintamumas), laikoma, kad taškai erdvėje išsidėstę grupėmis; o jei taškų skaičiaus gardelėse kintamumas nedidelis, laikoma, kad taškai erdvėje išsidėstę atsitiktinai. Šie skirtumai pavaizduoti 3.7 pav.



3.7 pav. Taškų erdvinio išsidėstymo analizė tinkleliu

Minėtieji kintamumo tipai 3.7a–3.7c pav. matyti gana aiškiai; bet (kaip ir praeitame skyriuje aptartu artimiausių kaimynų analizės atveju) jei iš šios analizės norime daryti kokias nors išvadas, taškų išsidėstymo tipą reikia nustatyti statistiniais metodais.

Nors analizės tinkleliu statistinis tikrinimas savo forma ir panašus į artimiausių kaimynų analizės tikrinimą, bet jį atlikti galima daug paprasčiau. Šiuo atveju reikia žinoti tik taškų skaičiaus gardelėse vidurkį ir dispersiją – ir vienas, ir kitas yra klasikinė aprašomoji statistika. Iš šių dviejų kintamųjų galima apskaičiuoti trečią statistiką:

$$VMR = \frac{\text{Dispersija}}{\text{Vidurkis}},$$

kur *VMR* yra dispersijos ir vidurkio santykio (*variance-mean ratio*) akronimas. Jei *VMR* lygus vienetui, taškų išsidėstymas erdvėje laikomas atsitiktiniu; jei *VMR* didesnis už vienetą, taškų išsidėstymas erdvėje laikomas labiau grupiniu negu atsitiktiniu; o jei *VMR* mažesnis už vienetą, taškų išsidėstymas erdvėje laikomas labiau tolygiu negu atsitiktiniu. Tikrinimo statistika yra chi kvadratas:

$$\chi^2 = VMR(m - 1),$$

kur *m* yra gardelių skaičius, o nulinė hipotezė – atsitiktinis erdvinis išsidėstymas. Kadangi šio pavyzdžio 3.7a paveikslėlyje taškų skaičius gardelėse dispersija lygi nuliui, statistiškai tikrinsime 3.7b ir 3.7c paveikslėlius. Reikalingi kintamieji surašyti 3.3 lentelėje.

3.3 lentelė. Erdvinių taškų išsidėstymo analizei tinkleliu reikalingi kintamieji

	3.7b pav.	3.7c pav.
Taškų skaičiaus gardelėse dispersija	2,5	21
Taškų skaičiaus gardelėse vidurkis	3	3
VMR	0,833	7
χ^2 statistika	6,67	56

Iš 3.3 lentelės matyti, kad 3.7b pav. turi aiškių atsitiktinio erdvinio išsidėstymo požymių, bet taip pat yra pagrindo manyti, kad jis daugiau tolygus, negu atsitiktinis. 3.7c pav. rezultatai rodo, kad yra aiškių grupių požymių, o chi kvadrato statistika aiškiai leidžia atmesti nulinę atsitiktinio erdvinio išsidėstymo hipotezę – kritinių reikšmių laisvės laipsnių skaičius yra septyni – devyni stebėjimai minus du skaičiuojami parametrai (vidurkis ir dispersija). Kadangi chi kvadrato statistika lygi 6,67, erdvinis taškų išsidėstymo skirtumas nuo atsitiktinio erdvinio išsidėstymo yra nereikšmingas. Taigi, visi 3.7 pav. pateikti išsidėstymo pavyzdžiai statistiniu požiūriu atitinka jų išvaizdą.

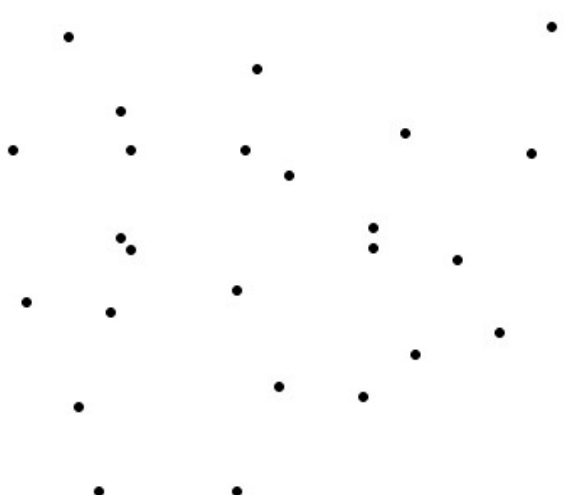
3.3.4 Erdvinio išsidėstymo analizės ribotumai

Kaip ir kiekviena erdvinė ar klasikinė statistinės analizės forma, erdvinio išsidėstymo analizė turi savų ribotumų. Pagrindinis ribotumas tas, kad ši analizė nepaaiškina, *kodėl* išsidėstymas yra vienoks, o ne kitoks. Užuoat aiškinusi priežastis, ji tiesiog nustato, koks yra išsidėstymas ir koks jis nėra.

Kita erdvinio pobūdžio problema susijusi su tiriamos srities plotu. 3.8 pav. pateikti du taškų išsidėstymo erdvėje pavyzdžiai. Taškų išsidėstymas 3.8a pav. atrodo atsitiktinis, o 3.8b pav. – grupinis. O reikalas tas, kad šiuose paveikslėliuose pavaizduotas tas pats išsidėstymas, tik skirtingais masteliais: 3.8a pav. žemėlapio mastelis stambesnis (mažesnis plotas) negu 3.8b pav.

[sivaizduokite, kad tiriate poveikį toksino, kurį kokia nors pramonės įmonė išleido į aplinkos orą tam tikro erdvėje pasiskirsčiusių taškų rinkinio centre. Jei analizuosite šio toksino įtaką atitinkamai ligai 3.8a pav. masteliu, toksiną gaminanti įmonė įtarimų nesukels. Bet jei sumažinsite analizės mastelį iki 3.8b pav. ir būtent šį mastelį laikysite tinkamu nagrinėjamai problemai, pramonės įmonė gali būti nurodyta kaip atitinkamos ligos priežastis.

Reikia žinoti, koks mastelis tinka analizei. Kartais tai gana lengva: mokslininkai gali būti nustatę, kokią toksino koncentraciją atmosferoje reikėtų laikyti pavojinga. Tada galima išmatuoti toksino koncentracijas įvairiais atstumais nuo pramonės įmonės ir pagal matavimų rezultatus apriboti nagrinėjamą sritį. Deja, ne visas panašios problemas taip „lengva“ išspręsti. Tokiu atveju apskritai teisingiausia būtų elgtis panašiai kaip su kintamų plotinių vienetų problema: nustatyti, kokie masteliai galėtų tikti, atlikti analizę šiais masteliais ir ieškoti rezultatų panašumo.

a) atsitiktinis?	b) grupinis?
	

3.8 pav. Mastelio problemos analizuojant erdvinį išsidėstymą

3.4 Kraštų įtaka

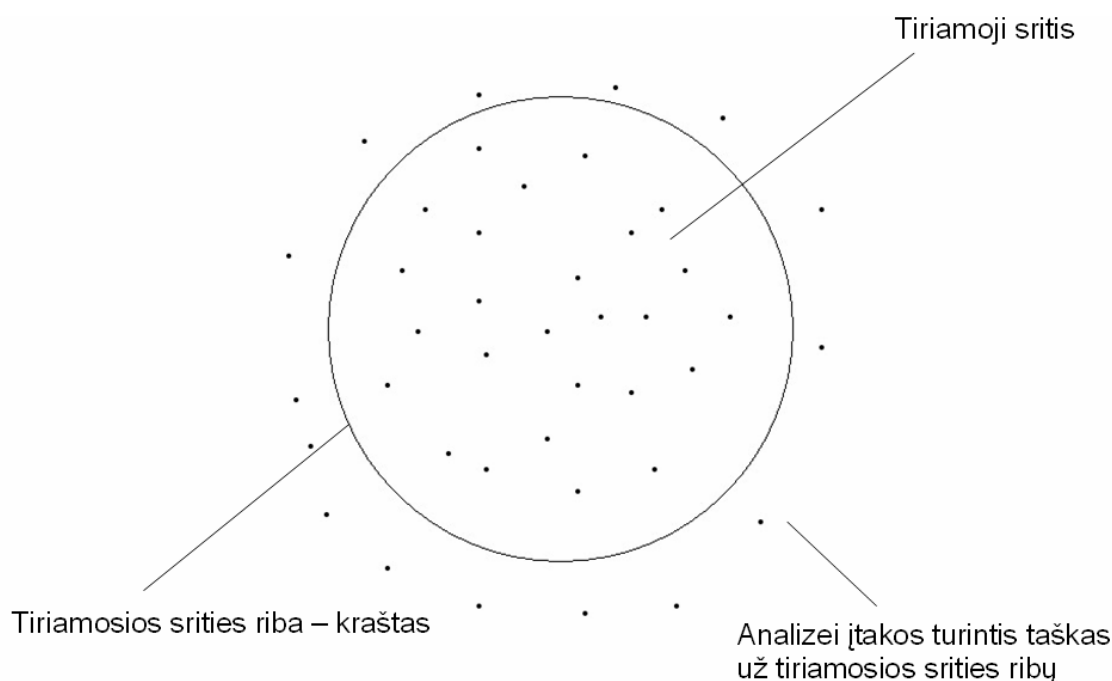
3.4.1 Kas yra kraštas ir kada jis turi įtakos?

Kraštas (*edge*) apibrėžiamas kaip nagrinėjamos erdvinės srities riba. Tai yra, išorinė visų analizuojamų erdvių vienetų riba. Pavyzdžiui, jei analizuojami erdviniai vienetai yra Lietuvos savivaldybės, kraštas yra Lietuvos valstybės siena.

Nors techniniu požiūriu bet kokios erdvinės analizės (ir bet kokios kitos analizės) sritis ribota, analizuojant taškų išsidėstymą kraštai yra ypatinga problema. Taip yra todėl, kad, kaip jau minėjome anksčiau, aprašomoji statistika skaičiuojama ir išvados daromos remiantis atstumais tarp įvykių arba tarp atsitiktinių taškų ir įvykių. Taigi, dėl kraštų įtakos kenčia atstumų skaičiavimu pagrįsti išsidėstymo analizės metodai.

Yra atvejų, kai kraštai problemų nekelia. Pavyzdžiui, dalis Lietuvos sienos yra riba su Baltijos jūra. Todėl vargu ar už šio krašto galėtų būti kokių nors įvykių – gali būti nedaug žmonių, gyvenančių visai šalia Lietuvos pakrantės, ir atitinkamų įvykių, bet atliekant analizę šiuos įvykius nesunku perkelti į sausumos sritį šalia jūros.

Problemų dėl kraštų atsiranda tada, kai tyrimų sritis yra dalis didesnio geografinio regiono, kuriame vyksta nagrinėjamas erdvinis procesas. Šiuo atveju už tiriamos srities ribų, bet didesnio geografinio regiono viduje vykstantys įvykiai gali turėti įtakos įvykiams tirmos srities viduje (Diglas, 2003). Tai ir vadinama kraštų įtaka. Jos pavyzdys pateiktas 3.9 pav.



3.9 pav. Kraštų įtaka

3.9 pav. pavaizduota tiriama sritis, jos ribos (kraštas) ir keli įvykiai už tiriamosios srities ribų, susiję su įvykiais tirmosios srities viduje. Tokia padėtis gali susidaryti tada, kai duomenis apie tam tikros ligos paplitimą vienoje savivaldybėje gauti įmanoma, o gretimose savivaldybėse – ne.

Puikus pavyzdys yra situacija, kai vienoje savivaldybėje (tiriamojame srityje) veikia atmosfera sklindančius teršalus į aplinką išleidžianti gamykla. Manoma, kad dėl didelio atstumo šios gamyklos išleidžiami teršalai nekenks gretimoms savivaldybėms, arba gretimos savivaldybės neprisideda prie atliekamo tyrimo, tarkim, dėl pavaldumo priežasčių. Bet kuriuo atveju, atliekant gamyklos taršos poveikio analizę tikrai reikėtų atsižvelgti ir į įvykius už tiriamosios srities ribų.

3.4.2 Galima kraštų įtaka analizei

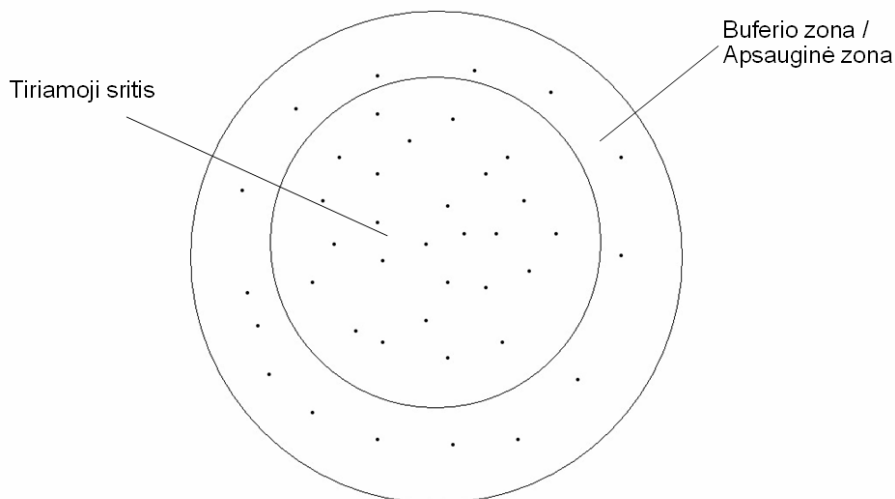
Kraštai gali iškreipti analizės rezultatus. Tai reiškia, kad rezultatai gali neatitikti tikrosios padėties. Dar blogiau – kadangi analizuojant trūko svarbios informacijos, rezultatai gali būti klaidinantys.

Derėtų pažymėti, kad ši situacija pasitaiko visose statistinės analizės srityse. Reikiamų duomenų dažnai nebūna, taigi neatsižvelgiama į svarbią informaciją. Tačiau, kaip ir kitose analizės srityse, yra metodų šiam duomenų trūkumui kompensuoti.

3.4.3 Kraštų įtakos korekcijos metodai

Kraštų įtaką galima įvertinti įvairiais metodais. Dažniausiai naudojami metodai yra išlyginimas (*adjustment*), apjuosimas žiedu (*toroidal wrapping*) ir buferio zona (*buffer zone*, kartais vadinama „apsaugine zona“ (*guard area*)) (Diggle, 2003).

Iš šių trijų kraštų įtakos įvertinimo metodų populiariausias yra buferio zonos metodas. Buferio zona sukuriamą tam tikru atstumu nuo išorinio tiriamosios srities krašto. Tai pavaizduota 3.10 pav.



3.10 pav. Kraštų įtakos korekcija kuriant buferio zoną

Nagrinėjant taškų išsidėstymą erdvėje, buferio zonoje esantiems taškams atstumai neskaičiuojami, bet į įvykius buferio zonoje leidžiama atsižvelgti atliekant su kitais įvykiais susijusius skaičiavimus (Bailey ir Gatrell, 1995). Iš esmės ši procedūra leidžia imituoti už tiriamosios srities ribų esančius taškus ir atlikti analizę nebijant iškreipti rezultatų. Taigi, naudojant buferio metodą, o taip pat ir išlyginimo metodą, nelieta rezultatų iškraipymo dėl kraštų įtakos, bet už tai sumokama padidėjusia sumažėjusių duomenų kiekio dispersija (Diggle, 2003).

Buferio zonos pločiui (atstumui nuo tiriamosios srities krašto) nustatyti standartinės taisyklės nėra. Galite atlikti analizę pasirinkę įvairaus pločio zonas – tuo pačiu atliksite ir jautrumo analizę.

3.5 Tankio skaičiavimas ir karštųjų zonų nustatymas

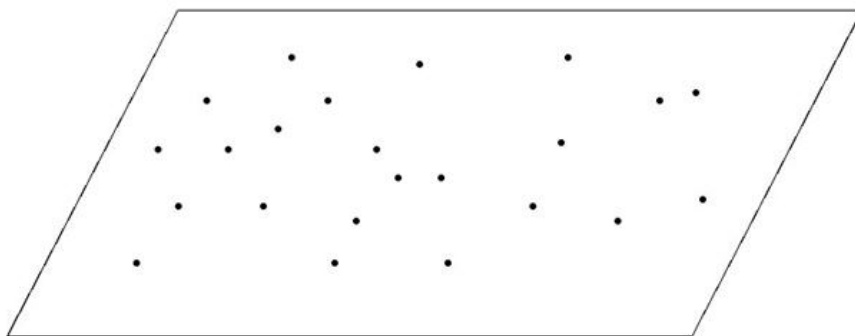
3.5.1 Tankio sąvoka

Vienas populiariausių metodų taškiniams duomenims vaizduoti tolydžiu paviršiumi yra branduolio tankio (*kernel density*) skaičiavimas. Paprastaiariant, branduolio tankio skaičiavimo metodas iš taškinių duomenų rinkinio įvykių sukuria gana sklandų paviršių, kurio reikšmės priklauso nuo taškinių įvykių tankio konkrečiose tiriamosios srities vietose. Šį tankį galima laikyti tiriamo reiškinio intensyvumu.

Viena iš branduolio tankio taikymo sričių yra karštųjų zonų (*hot spot*) nustatymas, paprastai naudojamas nusikalstamumui tirti. Užuoť bandžius išskirti išsidėstymo dėsninumus žemėlapyje su dešimtimis tūkstančių įvykių – nusikaltimų dažnai labai daug (pavyzdžiui, asmeninių daiktų arba automobilių vagysčių – žemėlapis tiesiog nusėtas taškais), skaičiuojamas branduolio tankis ir nustatomos „karščiausios“ vietos – vietos, kurioje taškų daugiausia. Šis erdvinės statistikos metodas lengvai įgyvendinamas ir patrauklus vizualiai.

3.5.2 Branduolio tankio skaičiavimas

Branduolio tankis skaičiuojamas trimis etapais. Įsivaizduokite 3.11 pav. pavaizduotą tiriamąją sritį: ji visa telpa į lygiagretainį; taškais lygiagretainio viduje pažymėtos įvykių tiriamojame srityje vietos; visos kitos tiriamosios srities vietos, įskaitant ir įvykių vietas, vadinamos taškais. Kitaip tariant, kaip jau minėjome anksčiau, taškas yra bet kuri tiriamosios srities vieta, o įvykis – vieta, kurioje kažkas įvyko.

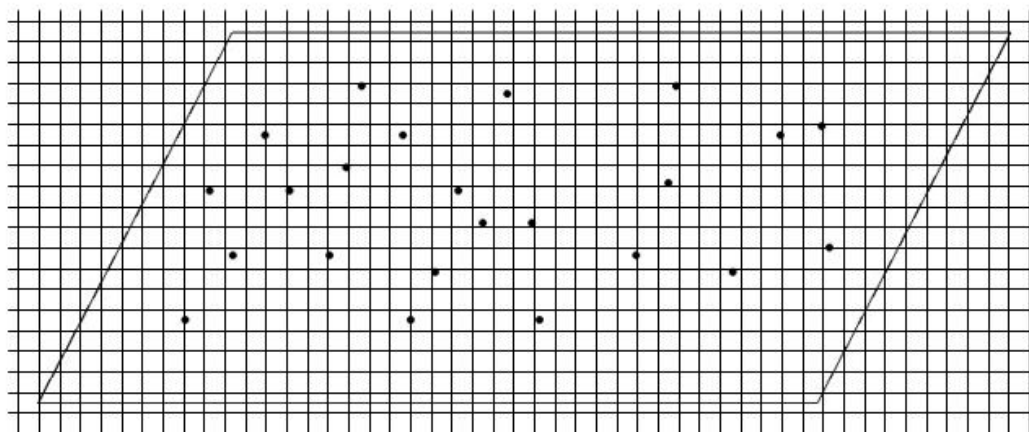


3.11 pav. Branduolio tankio skaičiavimo sritis

Pirmuoju branduolio tankio skaičiavimo etapu visa tiriamoji sritis uždengiama smulkiu stačiakampiu tinkleliu, net jei kai kuriose vietose įvykių ir nėra (3.12 pav.). Viena gera branduolio tankio skaičiavimo savybė yra ta, kad galima pačiam pasirinkti tinklelio gardelės dydį. Gera ji todėl, kad greičiausiai gardelės dydis priklausys nuo konkrečios užduoties, o keičiant gardelės dydį galima atlikti ir jautrumo analizę.

Jautrumo analizė gali būti ypač svarbi tose taikomosiuose srityse, kuriose nėra sutarta dėl gardelės dydžio – žinomas tik intervalas, kuriame gardelių dydį rinkosi kiti žmonės. Tokiu

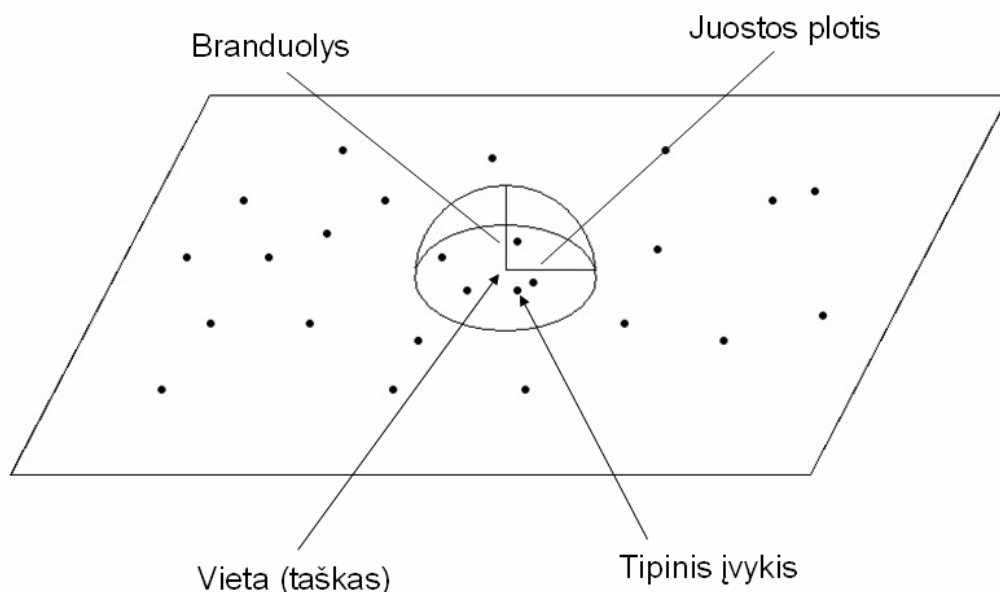
atveju galima apskaičiuoti branduolio tankį įvairiais gardelės dydžiais iš šio intervalo ir nustatyti, ar rezultatai priklauso nuo gardelės dydžio. Retklifas (1999) taip pat pasiūlė metodą tinkamo gardelės dydžiui nustatyti tuo atveju, kai nėra kitų metodų: aplink tiriamąją sritį nubrėžti mažiausią įmanomą stačiakampį, išmatuoti trumpąją jo kraštinę (metrais) ir padalinti jos ilgį iš šimto penkiasdešimties.



3.12 pav. Branduolio tankio tinklelis

Antruoju etapu kiekvienai tiriamosios srities gardelei skaičiuojama tankio (intensyvumo) funkcijos reikšmė. Nustatoma kiekvienos gardelės centro padėtis – centroidas, aplink kurį brėžiamas nurodyto spindulio apskritimas. Šio apskritimo spindulys vadinamas branduolio tankio skaičiavimo juostos plotiu. Kuo daugiau įvykių patenka į šį apskritimą, tuo didesnė šios gardelės tankio reikšmė. Be to, skaičiuojant tankį įvykiams suteikiami skirtingi svoriai: įvykio svoris priklauso nuo jo atstumo iki centroido, todėl arčiau centroido esančių įvykių svoris didesnis, taigi proporcingai didesnis ir indėlis į galutinę gardelės tankio reikšmę. Tai pavaizduota 3.13 pav.

Svarbiausia – nustatyti tinkamą spindulį (juostos plotį), nes ši branduolio tankio metodo dalis yra jautriausia pokyčiams. Kaip ir anksčiau aptartas gardelės dydis, juostos plotis priklauso nuo konkrečios užduoties. Galima pažiūrėti, koks juostos plotis buvo naudotas panašioms tyrimams, arba sužinoti tinkamą intervalą ir atlikti jautrumo analizę.



3.13 pav. Branduolio tankio skaičiavimas

Trečiuoju etapu branduolio tankio skaičiavimo rezultatai pavaizduojami žemėlapyje. Kadangi apskaičiuotos reikšmės atitinka tam tikrą reikšmių intervalą, geriausia rezultatus vaizduoti graduotų spalvų palete: pavyzdžiui, balta arba rausva spalva – mažas tankis, tamsiai raudona – didelis.

Dar branduolio tankio metodas patogus tuo, kad gardelių tankis išreiškiamas vienetais, turinčiais prasmę nagrinėjant taškų išsidėstymą erdvėje – pavyzdžiui, įvykių skaičiumi kvadratiniam kilometrui. Iš to išplaukia, kad gardelių reikšmių palyginimas taip pat turi prasmę: jei 122-os tinklelio gardelės reikšmė yra 100, o 123-ios – 50, tai 122-oje gardelėje įvykiai yra dukart intensyvesni negu 123-oje.

Panagrinėkime kelis paprastus trumpus pavyzdžius. Pirmiausia reikia žinoti branduolio tankio formulę. Ši formulė aprašo ketvirto laipsnio branduolį:

$$\lambda = \sum_{d_i \leq r} \left(\frac{3}{\pi r^2} \right) \left(1 - \frac{d_i^2}{r^2} \right)^2,$$

kur λ yra intensyvumo reikšmė (branduolys), $\pi = 3,14$, r – apskritimo spindulys (juostos plotis), o d_i – atstumas tarp centroido ir į juostos plotį patenkančio įvykio i . Pasirinkus juostos plotį, $\frac{3}{\pi r^2}$ tampa konstanta, kurios šiuose pavyzdžiuose galime nepaisyti. Lieka tik $\left(1 - \frac{d_i^2}{r^2} \right)^2$, o kad būtų dar paprasčiau, skliausteliuose esančios išraiškos nekelkime kvadratu. Tegu $r = 10$. Dabar įsivaizduokime du galimus atvejus.

Pirmuoju atveju į juostos plotį patenka keturi taškai: du iš jų nuo centroido nutolę dviem vienetais, kiti du – trim vienetais. Antru atveju į juostos plotį patenka du taškai, abu nutolę nuo centroido devyniais vienetais. Apskaičiuokime tankį abiem atvejams.

Pirmuoju atveju λ reikšmė yra 3,74, antruoju – 0,38; aišku, jei skaičiuotume pagal visą formulę, gautume kitas reikšmes.

$$\lambda = \left(1 - \frac{4}{100}\right) + \left(1 - \frac{4}{100}\right) + \left(1 - \frac{9}{100}\right) + \left(1 - \frac{9}{100}\right)$$

1 atvejis:

$$= \frac{96}{100} + \frac{96}{100} + \frac{91}{100} + \frac{91}{100}$$

$$= \frac{374}{100} = 3.74$$

$$\lambda = \left(1 - \frac{81}{100}\right) + \left(1 - \frac{81}{100}\right)$$

2 atvejis:

$$= \frac{19}{100} + \frac{19}{100}$$

$$= \frac{38}{100} = 0.38$$

Gauta λ reikšmė pirmuoju atveju daug didesnė dėl dviejų priežasčių. Pirmą, į juostos plotį patenka daugiau įvykių, antra, pirmu atveju į juostos plotį patenkantys įvykiai yra arčiau centro. λ reikšmė priklauso nuo šių abiejų veiksnių, todėl įmanoma, kad vienos tinklės gardelės reikšmė bus didesnė už kitos net ir tada, kai į antrosios gardelės juostos plotį pateks daugiau įvykių. Trivialus pavyzdys: jei į juostos plotį nepatenka nė vienas įvykis, $\lambda = 0$.

3.5.3 Tankio skaičiavimo apribojimai

Branduolio tankio skaičiavimo taikymą daugiausia riboja du dalykai: metodo teisingumas ir įvykių vaizdavimo tikslumas.

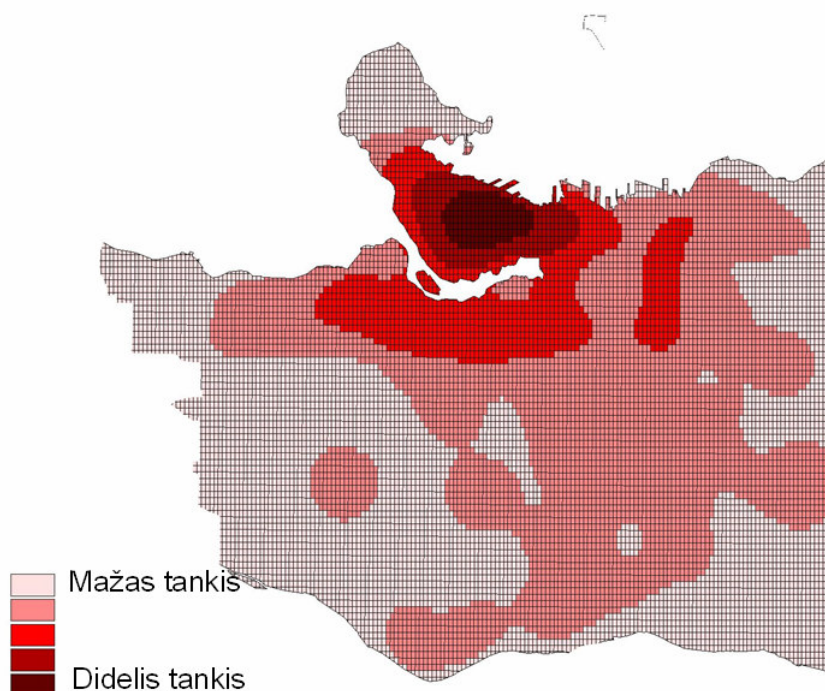
Abejonės dėl metodo teisingumo turbūt yra „griežčiausia“ branduolio tankio skaičiavimo kritika. Taškiniai duomenys būna dviejų tipų. Pirmojo tipo taškiniai duomenys vaizduoja diskrečius įvykius, pavyzdžiui, atskirus nusikaltimus, eismo įvykius ar ligos atvejus. Antrojo tipo taškiniai duomenys vaizduoja tolydaus paviršiaus matavimus, pavyzdžiui, temperatūrą ar užterštumą. Kadangi antrojo tipo duomenys iš tikro vaizduoja tolydų paviršių, yra prasmės prieš atliekant analizę transformuoti juos į tolydų paviršių (išmatuoti, pavyzdžiui, temperatūrą visuose šalies taškuose būtų neįmanoma vien dėl tokio matavimo kainos).

Tačiau pirmojo tipo taškiniais duomenimis vaizduojami reiškiniai nėra tolydūs. Todėl daug žmonių teigia, kad nėra prasmės transformuoti taškinius įvykius į tolydų paviršių, kadangi patys duomenys yra diskretūs. Diskrečius įvykius vertindami tolydziais paviršiais tam tikra prasme „nusikalsta“ asmenys, sudarantys nusikalstamumo žemėlapius.

Bet ši kritika nebus tokia problemiška, jei laikysime, kad tolygus paviršius vaizduoja riziką. Gali būti, kad nusikaltimas visai atsitiktinai įvyko konkrečiame taške, o ne kitame taške dviem gardelėmis toliau, taigi rizika tapti jo auka abiejose gardelėse gali būti vienoda.

Sunkiau su tais taškais (ir mažomis aplinkinėmis sritimis), kur niekada neįvyko jokių nusikaltimų, bet branduolio tankio funkcija ir šioms vietoms priskiria tam tikrą nusikaltimo riziką. Taigi, kaip ir kiekvienos statistinės analizės atveju, branduolio tankio skaičiavimo rezultatus interpretuoti reikia atsargiai.

Su antruoju apribojimu susitvarkyti nesudėtinga. Jis atsiranda dėl to, kad daugelis tyrinėtojų skaičiuodami branduolio tankį atsižvelgia tik į įvykių duomenis. Gauti rezultatai tikrai nurodo didžiausio aktyvumo zonas, tačiau ne visada teisingose vietose.

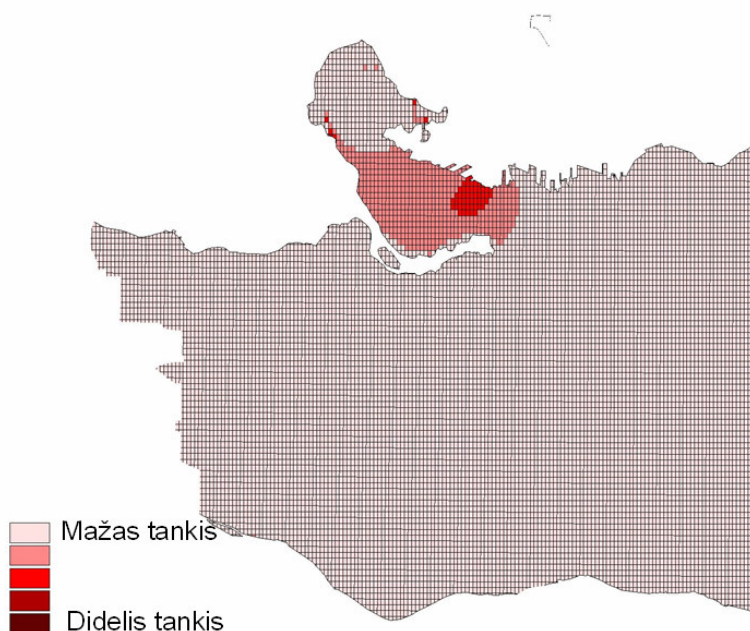


3.14 pav. Viengubo branduolio tankis

3.14 pav. pavaizduotas branduolio tankio žemėlapis, gautas tik iš įvykių duomenų – skaičiavimams naudotas tik vienas taškinis duomenų rinkinys. Aiškiai matyti karštoji zona – didelio tankio sritis netoli pusiasalio centro žemėlapio viršuje (šis pavyzdys ne Lietuvos). Bet būtina atsakyti į vieną klausimą: ar pati ši vieta kuo nors ypatinga, ar karštoji zona nustatyta tik todėl, kad ten didelis konkrečios rizikos grupės gyventojų tankis?

Nusikalstamumas, sergamumas, eismo įvykių dažnis ir pan. skaičiuojami tam tikram gyventojų skaičiui, nes norint tiksliai įvertinti šią riziką, reikia atsižvelgti į nusikaltimo, ligos ar eismo įvykių rizikos grupės dydį. Daugiau gyventojų turinčiose savivaldybėse nebūtinai padaroma daugiau nusikaltimų, įvyksta daugiau avarių ar daugiau žmonių serga. Svarbiau žinoti, ar daugiau nusikaltimų, eismo įvykių ir ligos atvejų tenka vienam rizikos grupės atstovui (paprastai rizikos grupės dydis lygus tiriamosios srities gyventojų skaičiui, eismo įvykių atveju – eismo apimčiai). Taigi, ar 3.14 pav. nurodytas tikrai didelis aktyvumas, ar toje vietoje tiesiog daugiau žmonių?

Šio apribojimo išvengiama vadinamuoju dvilypio branduolio tankio (*dual kernel density*) metodu. Norint apskaičiuoti dvilypio branduolio tankį, reikia dviejų duomenų rinkinių: analizuojamų įvykių duomenų ir šių įvykių rizikos grupės duomenų. Skirtumas tarp gautų žemėlapių gali būti stulbinantis. Jis matyti iš 3.14 ir 3.15 pav.



3.15 pav. Dvilypio branduolio tankis

Dvilypio branduolio tankio žemėlapyje (3.15 pav.) tokio didelio tankio, koks pavaizduotas 3.14 pav., toli gražu nėra. Didžiausio tankio vietos abiejuose žemėlapiuose beveik sutampa, bet reiškinių vaizdas gerokai skiriasi. Taigi, gavę bet kokį tankio žemėlapi, pirmiausia paklauskite: ar tai vienalypio, ar dvilypio branduolio tankio žemėlapis? Jei atsakymas bus „vienalypio“ arba „nežinau“, tada rezultatus interpretuoti reikia atsargiai, nes žemėlapyje neatsižvelgta į rizikos grupės dydį.

3.6 Vietinė erdvinė statistika

3.6.1 Globalios erdvinės statistikos apribojimai

Visas mūsų iki šiol nagrinėtas erdvinės statistikas galima pavadinti „globaliomis“ (*global spatial statistics*). Iš tiesų dar maždaug prieš dvidešimt metų taip buvo galima vadinti absoliučiai visą erdvinę statistiką. „Globali“ reiškia, kad iš viso duomenų rinkinio gaunamas vienas statistinių rezultatų rinkinys, kurio paskirtis – aprašyti visą tiriamą sritį. Šis statistinių rezultatų rinkinys yra vidurkis pagal visus analizuojamus erdvinius vienetus. Tačiau jei analizuojami reiškiniai tiriamojoje srityje stipriai kinta, globalios erdvinės statistikos vertė labai menka. Ekstremalus bevertės globalios statistikos pavyzdys yra vidutinė temperatūra Rusijoje. Rusija tokia didelė, kad viena statistika apie visus regionus neteikia praktiškai jokios naudingos informacijos. Panagrinėkime kitą pavyzdį.

Tarkime, kad suskaičiavome automobilių vagysčių Morano I statistiką Lietuvos savivaldybėse. Gautume tokį rezultatą: Morano $I = 0,02$. Tai reiškia, kad autokoreliacija erdvėje teigiama, bet labai maža ir statistiškai nereikšminga. Tačiau jei kiekvienai iš šešiasdešimties Lietuvos savivaldybių apskaičiuotumėte vietinę Morano I statistikos versiją (išsamiai aprašytą toliau), pamatytumėte, kad vietiniai Morano I rezultatai kinta nuo $-4,89$ iki $3,06$, o jų vidurkis yra $0,106$. Be to, penkios iš apskaičiuotųjų vietinių Morano I statistikos reikšmių labai skiriasi nuo nulio, priešingai negu globalus Morano I rezultatas.

Kadangi šie rezultatų skirtumai panašūs daugumoje tyrimų sričių, neseniai erdvinės statistikos moksle atsirado kryptis, kurios atstovai stengiasi aptikti statistinių rezultatų skirtumus erdvėje, užuot darę prielaidą, kad šių skirtumų nėra. Iš tiesų kiekybinei geografijai seniai buvo priekaištaujama, kad ji ieško globalių bendrybių ir nesistengia ieškoti vietinių. Ši erdvinės statistikos, plačiąja prasme – erdvinės analizės kryptis vadinama vietine analize, vietiniu modeliavimu arba vietine erdvine statistika.

3.6.2 Vietinės erdvinės statistikos nauda

Daugeliu atvejų vietinė erdvinė statistika (*local spatial statistics*) duoda rezultatų, kurių neįmanoma gauti globalios erdvinės statistikos metodais – bet reikia pažymėti, kad vietinę erdvinę statistiką geriausia naudoti kartu su globalia erdvine statistika, o ne vietoje jos. Iš pateikto pavyzdžio matyti, kad vietinė erdvinė statistika daugiareikšmė, kinta erdvėje ir išryškina erdvinius skirtumus visame tiriamame regione.

Aišku, kad tai stiprus argumentas už vietinės statistikos naudojimą, bet analizei ji naudinga ir dėl kitų priežasčių. Vietinę erdvinę statistiką galima pavaizduoti žemėlapyje. Kaip jau minėjome nagrinėdami Lietuvos savivaldybių pavyzdį, kiekvienas plotinis erdvinis vienetas turi savo vietinę erdvinę statistiką. Todėl ši statistika yra tiesiog dar vienas duomenų rinkinio kintamasis, kurį galima pavaizduoti žemėlapyje. Dėl šios savybės vietinė statistika laikoma patogia dirbant su GIS. Ir pagaliau vietinę erdvinę statistiką galima panaudoti ieškant karštųjų zonų: įsivaizduokite du arba daugiau erdvinių vienetų, kuriuose vietinė autokoreliacija erdvėje teigiama ir statistiškai reikšminga.

Pirmasis plačiai žinomas vietinės erdvinės analizės pasiekimas susijęs su erdviniu taškų išsidėstymu – tai Oupenšo ir bendraautorų 1987 m. sukurta geografinės analizės mašina (*geographical analysis machine*). Ši geografinės analizės mašina – tiesa, išplėsta daugelio kitų mokslininkų – yra daug skaičiavimų atliekantis algoritmas, į kurio smulkmenas dabar

visai nesigilinsime. Tačiau jo bendra formuluotė verta dėmesio, nes ji kartojasi daugumoje vietinės erdvinės statistinės analizės sričių.

Pirma, yra metodas tiriamajai sričiai sudalinti į smulkesnius vienetus; antra, yra metodas taškų išsidėstymui kiekviename iš šių smulkesnių vienetų aprašyti; trečia, yra metodika, pagal kurią nustatomi nuo vidurkio esminiai besiskiriantys taškų išsidėstymai; ir ketvirta, yra metodas rezultatams vaizduoti žemėlapyje. Nors geografinės analizės mašina gana smarkiai skiriasi nuo kai kurių naujausių vietinės erdvinės statistinės analizės metodikų, minėtieji keturi elementai vienokia ar kitokia forma aptinkami visuose vietinės erdvinės statistikos metoduose.

3.6.3 Vietinė Morano I

Anselinas (1995) įvedė vietinę globalios Morano I atmainą, kurią pavadino vietine Morano I. Apskritai, Anselinas (1995) aprašė visą vietinių erdvinių statistikų klasę – vadinamuosius vietinius erdvinių ryšių rodiklius (*local indicators of spatial association* – LISA). Šie vietiniai erdvinio ryšio rodikliai turi dvi savybes, išskiriančias juos iš kitų vietinių erdvinių statistinių įverčių: vietinis erdvinio ryšio rodiklis turi nurodyti panašių reikšmių klasterių buvimo aplink analizuojamą erdvinį vienetą laipsnį, o visų vietinių erdvinio ryšio rodiklių suma turi būti proporcinga atitinkamam globaliam rodikliui – šiuo atveju globaliai Morano I. Kitaip tariant, vietiniai ir globalus erdvinių ryšių rodikliai turi būti susiję, priešingu atveju vietinių ir globalių statistikų pavadinimai neturi būti panašūs, kad nekiltų painiavos.

Vietinė Morano I skaičiuojama pagal formulę:

$$\text{Vietinė Morano } I = \frac{(y_i - \bar{y}) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_j - \bar{y})}{\left(\sum_{i=1}^n (y_i - \bar{y})^2 \right) / n}$$

kurios visi kintamieji apibrėžti skyriuje apie autokoreliaciją erdvėje. Atkreipkite dėmesį, kad vietinė Morano I versija labai panaši į globalią Morano I, o tai leidžia šį erdvinį statistinį įvertį priskirti vietiniams erdvinio ryšio rodikliams. Nors ši formulė sudėtinga – kaip ir globalaus erdvinio ryšio rodiklio, ArcGIS aplinkoje Morano I lengva apskaičiuoti ArcToolbox programa.

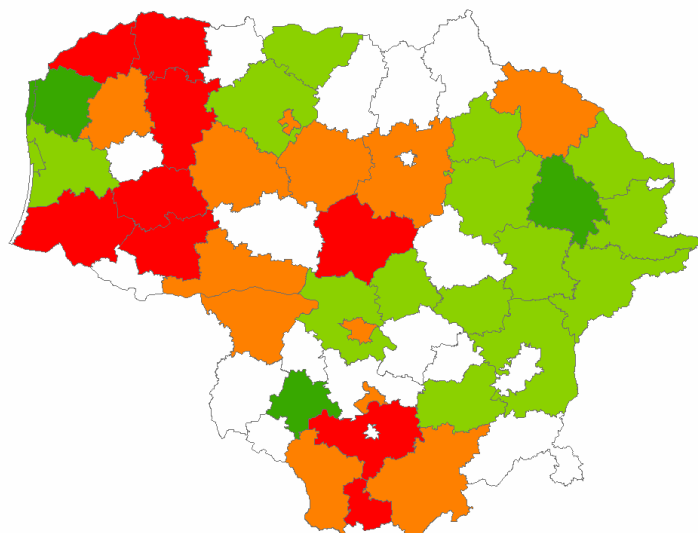
3.6.4 Vietinė klasterių analizė

Vietinė klasterių analizė (*local cluster analysis*), kaip galima spėti iš paties pavadinimo, ieško erdvinių objektų klasterių vietiniu lygmeniu. Priešingai negu aptarta 3.3 skyriuje, čia klasteriai sudaromi iš plotinių erdvinių vienetų. Šiuo konkrečiu atveju ieškoma vietinių klasterių, sudarytų iš savivaldybių su reikšmingomis vietinėmis Morano I reikšmėmis.

Vietinė klasterių analizė pagal vietinę Morano I statistiką atliekama keliais etapais. Pirmiausia reikia apskaičiuoti vietinę Morano I statistiką. Atlikus šį veiksmą, ArcGIS programoje atsiranda dviejų naujų kintamųjų stulpeliai: LMiContig ir LMzContig. LMiContig yra vietinės Morano I statistikos reikšmė, o LMzContig – su kiekviena LMi reikšme susijusi z-statistika; „Contig“ (*contiguity* – gretimumas) reiškia, kad naudojama anksčiau minėta gretimumo erdvinių svorių matrica.

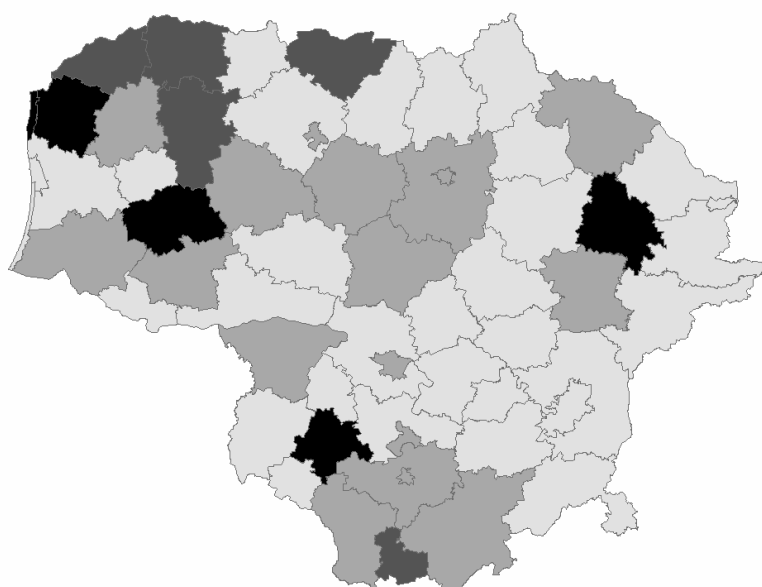
Pirmas kintamasis, LMiContig, pavaizduotas 3.16 pav. Žalios spalvos atspalviai (nuo šviesaus iki tamsaus) vaizduoja į neigiamą pusę didėjančias Lietuvos savivaldybių vietinės Morano I statistikos reikšmes, o raudonos (nuo šviesaus iki tamsaus) – didėjančias

teigiamas vietinės Morano / statistikos reikšmės. Baltos savivaldybės – tai erdviniai vienetai, kurių vietinė Morano / statistika artima nuliui. Kaip ir galima buvo tikėtis, Lietuvos savivaldybių vietinės Morano / statistikos kintamumas gana didelis, ir daug vietinės Morano / statistikos reikšmių yra neigiamos – priešingai negu anksčiau gauta statistiškai nereikšminga globali autokoreliacija erdvėje. Nepaisant šio geografinio kintamumo, reikia nustatyti šios vietinės Morano / statistikos reikšmingumą bei visų gautų vietinių klasterių pobūdį.



3.16 pav. Automobilių vagysčių skaičius 10 000 gyventojų: vietinė Morano I

Vietinės Morano / statistikos reikšmingumas pavaizduotas 3.17 pav. (tamsesnis atspalvis – didesnis reikšmingumas); statistiškai reikšmingos tik visiškai juodos sritys. Žiūrint į šiuos du žemėlapius ir žinant, kad globali autokoreliacija erdvėje yra statistiškai nereikšminga, atsiveria visiškai kitoks autokoreliacijos Lietuvos savivaldybėse vaizdas. Nors vietinės Morano / statistikos reikšmių kintamumas didelis, dauguma jų yra statistiškai nereikšmingos – tai atitinka globalią Morano / statistiką. Kitas etapas – atrinkti tas Lietuvos savivaldybes, kurių vietinė Morano / statistika yra statistiškai reikšminga, nustatyti šių savivaldybių pobūdį ir jas pavaizduoti žemėlapyje.



3.17 pav. Vietinės Morano / statistinis reikšmingumas (Z-įvertis)

Galutinis šio pavyzdžio žemėlapis pateiktas 3.18 pav. Šiame žemėlapyje pavaizduotos Lietuvos savivaldybės, kurių vietinė Morano / statistika reikšminga 10 procentų (absoliuti z-įverčio reikšmė $> 1,28$). Statistiškai reikšminga yra penkių Lietuvos savivaldybių vietinė Morano / statistika.

Žemėlapyje vaizduojamos šios kategorijos: Didelė–Maža (rausva), Maža–Didelė (mėlyna) ir Maža–Maža (tamsiai mėlyna); dar gali būti ir Didelė–Didelė kategorija, bet jos šiame žemėlapyje nėra. Didelė–Didelė ir Maža–Maža kategorijų savivaldybių vietinė autokoreliacija erdvėje yra teigiama (vietinės Morano / reikšmės panašios), Didelė–Maža ir Maža–Didelė savivaldybių – neigiama (vietinės Morano / reikšmės nepanašios).

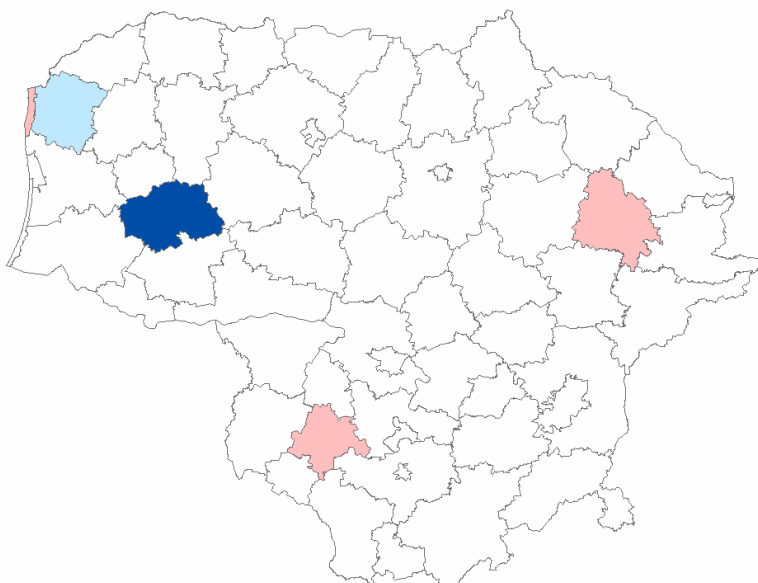
Ketrios įmanomos savivaldybių kategorijos apibūdinamos taip:

Didelė–Didelė – statistiškai reikšminga teigiama vietinė Morano I statistika, didelis automobilių vagiamumas, aplinkinėse savivaldybėse automobilių vagiamumas taip pat didelis;

Didelė–Maža – statistiškai reikšminga neigiama vietinė Morano I statistika, didelis automobilių vagiamumas, aplinkinėse savivaldybėse automobilių vagiamumas mažas;

Maža–Didelė – statistiškai reikšminga neigiama vietinė Morano I statistika, mažas automobilių vagiamumas, aplinkinėse savivaldybėse automobilių vagiamumas didelis;

Maža–Maža – statistiškai reikšminga teigiama vietinė Morano I statistika, mažas automobilių vagiamumas, aplinkinėse savivaldybėse automobilių vagiamumas taip pat mažas.



3.18 pav. Klasterių su reikšminga vietine Morano / statistika tipai

Nors šiuo atveju iš šešiasdešimties Lietuvos savivaldybių statistiškai reikšmingos vietinės Morano / reikšmės gautos tik penkiose, vietinės statistikos nauda turėtų būti akivaizdi. Jei atsižvelgtume tik į globalų autokoreliacijos erdvėje matą, nuspręstume, kad jokių Lietuvos savivaldybių klasterių pagal automobilių vagiamumą nėra. Tačiau ištyrę vietinę autokoreliaciją, nesunkiai pastebėsime sritis, kur autokoreliacija reikšminga. Šia analize negalime atsakyti į klausimus, kokios yra konkrečios šių aukšto ar žemo automobilių vagiamumo klasterių atsiradimo priežastys, tačiau, jei būtume naudoję tik globalius autokoreliacijos erdvėje rodiklius, tokių klausimų net negalėtume užduoti.

Literatūra

1. Anselin, L. (1995) Local indicators of spatial association – LISA. *Geographical Analysis* 27: 93 – 115 p.
2. Bailey, T.C. and A.C. Gatrell (1995) *Interactive Spatial Data Analysis*. Harlow, England: Prentice Hall.
3. Brantingham, P.L. and P.J. Brantingham (1995) Location quotients and crime hot spots in the city. In C.R. Block, M. Dabdou, S. Fregly (eds.) *Crime Analysis through Computer Mapping*. Washington, DC: Police Executive Research Forum, 129 – 149 p.
4. Brantingham, P.L. and P.J. Brantingham (1998) Mapping crime for analytic purposes: location quotients, counts and rates. D. Weisburd and T. McEwen (red.) *Crime Mapping and Crime Prevention*. Monsey, NY: Criminal Justice Press, 263 – 288 p.
5. Diggle, P.J. (2003) *Statistical Analysis of Spatial Point Patterns, Second Edition*. London: Arnold Publishers.
6. Isard, W., I.J. Azis, M.P. Drennan, R.E. Miller, S. Saltzman, and E. Thorbecke (1998) *Methods of Interregional and Regional Analysis*. Brookfield, VT: Ashgate Publishing Limited.
7. Miller, M.M., L.J. Gibson, and N.G. Wright (1991) Location quotient: a basic tool for economic development studies. *Economic Development Review* 9: 65 – 68 p.
8. Openshaw, S. (1984) *The Modifiable Areal Unit Problem*. CATMOG (Concepts and Techniques in Modern Geography) 38. Norwich: Geo Books.
9. Openshaw, S., M. Charlton, C. Wymer, and A. Craft (1987) Developing a mark 1 geographical analysis machine for the automated analysis of point data. *International Journal of Geographical Information Systems* 1: 335 – 358 p.
10. Ratcliffe, J. (1999) Hotspot Detective for MapInfo Helpfile Version 1.0.
11. <http://jratcliffe.net/ware/index.htm>
12. Robinson, W.S. (1950) Ecological correlations and the behaviour of individuals. *American Sociological Review* 15: 351 – 357.
13. Tobler, W. R. (1970) A computer model simulation of urban growth in the Detroit region. *Economic Geography* 46: 234 – 240 p.

Papildoma informacija

Klasikinė koreliacija

- The Statistics Homepage:
 - <http://www.statsoft.com/textbook/stathome.html>
- Introductory Statistics:
 - <http://www.psychstat.missouristate.edu/sbk00.htm>
- HyperStat Online:
 - <http://davidmlane.com/hyperstat/>

Kintamo plotinio vieneto problema ir ekologinė klaida

- Džerio Retklifo svetainė (KPVP):
 - <http://www.jratcliffe.net/research/maup.htm>
- Džerio Retklifo svetainė (ekologinė klaida):
 - <http://www.jratcliffe.net/research/ecolfallacy.htm>

Tankio skaičiavimas ir karštųjų zonų nustatymas

- Mapping Crime: Understanding Hot Spots (U.S. National Institute of Justice):
 - <http://www.ojp.usdoj.gov/nij/maps/ncj209393.html>
- CrimeStat Manual (8 skyrius):
 - <http://www.icpsr.umich.edu/CRIMESTAT/files/CrimeStatChapter.8.pdf>
- arba žiūrėkite turinyje (*Table of Contents*) adresu:
 - <http://www.icpsr.umich.edu/CRIMESTAT/download.html>

4. Geostatistika

Geostatistika – vienas iš sudėtingiausių erdvinės analizės aspektų. Ji apima paviršiaus reikšmių interpoliacijos, glodinimo, vertinimo ir prognozės pagal diskrečiuosius matavimus nagrinėjimą. Šioje kurso dalyje pateikiami geostatistikos, įskaitant tiriamosios erdvinės duomenų analizės elementus, struktūrinės analizės, įskaitant gretimų paviršių ypatybių apskaičiavimą ir modeliavimą, ir paviršiaus prognozės bei rezultatų vertinimo pagrindai. Aptarti metodai: atvirkščiai proporcingo atstumo metodas (*Inverse Distance Weighting*), trendo analizė, naudojant globalius ir lokalius polinomus, interpoliacija splineais ir krigingo prognozių technikos. Nagrinėjamos ir priklausomybės bei autokoreliacijos erdvėje sąvokos. Aptariamas geostatistinių metodų skirstymas į kategorijas, šių metodų taikymo rekomendacijos ir rezultatų įvertinimo modeliai.

Dalies turinys

- Geostatistikos pagrindai
- Geostatistinių metodų klasifikacija: globalūs ir lokalūs, tikslūs ir netikslūs, determinuotieji ir tikimybiniai
- Interpoliacija ir glodinimo metodai:
- Krigingo (*Kriging*) prognozavimas
- Modelio įvertinimas
- Išvados. Geostatistinių metodų palyginimas

4.1 Geostatistikos pagrindai

4.1.1 Motyvacija

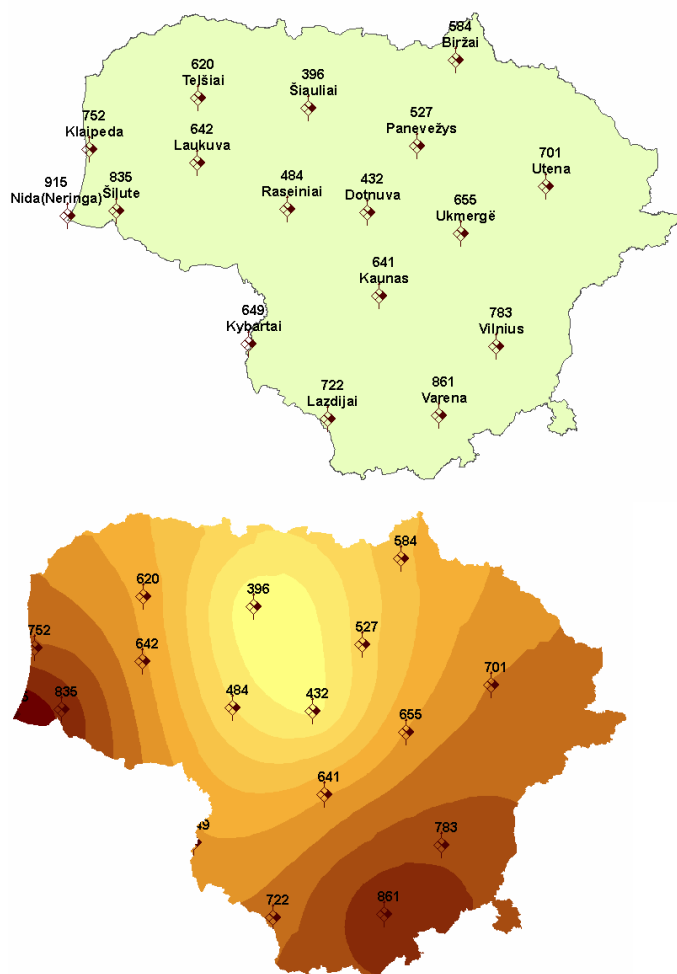
Dažnai reikia tirti natūralių ar socialinių darinių erdvinės struktūros, naudojant taškų stebėjimą ir kiekybinę analizę. Pvz., rūgščius kritulius – pagrindinę miškų nykimo priežastį – stebi atskiros stotys. Terminas „[rūgštūs lietūs](#)“ reiškia atmosferos kritulių (lietaus ar sniego) užteršimą rūgštimis. Medžiagų, ištirpusių vandenyje, rūgštingumas matuojamas [pH](#) (apibrėžiamas kaip neigiamas vandenilio jonų koncentracijos logaritmas). Pagal šią matavimo skalę tirpalai, kurių pH yra mažesnis nei 7, vadinami [rūgštiniais](#). Jei pH yra didesnis nei 7,0, tirpalas yra [šarminis](#). Dėl natūralių atmosferinių reakcijų, kuriose dalyvauja anglies dioksidas, įprastinis kritulių pH yra nuo 5,0 iki 5,6. Palyginimui, distiliuoto vandens, kuriame nėra jokių kitų medžiagų, pH yra 7,0. Krituliai laikomi rūgščiais, kai pH yra mažesnis nei 5,6, t.y. 25 kartūs rūgštesni nei grynas distiliuotas vanduo.

Kritulių prognozės gali būti pagrįstos ribotoje erdvėje paskirstytų meteorologinių stočių duomenimis. Tokio metodo apribojimai – didelės duomenų rinkimo išlaidos ir kai kurių regionų nepasiekiamumas. Pvz., Lietuvoje krituliams ir kitiems klimato parametrams stebėti buvo sukurtas 16 stočių meteorologinis tinklas. Kasmetinis kritulių kiekis nuo 2005 m. svyruoja nuo $915 \geq Pr \geq 396$ mm (žr. 4.1 lentelę).

4.1 lentelė. kritulių lygis 2005 m. ir stebėjimo stočių koordinatės

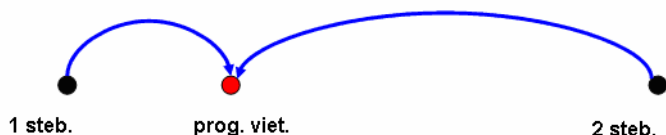
PAVADINIMAS	ID	mm	X	Y
Biržai	1	584	546484	6229692
Telšiai	13	620	389573	6206379
Utena	15	701	601070	6152592
Raseiniai	10	484	443915	6138743
Šilutė	12	835	339702	6137606
Dotnuva	2	432	492560	6136744
Nida(Neringa)	8	915	310049	6134547
Ukmergė	14	655	549330	6123740
Kybartai	4	649	420364	6056681
Varėna	16	861	535976	6012783
Lazdijai	7	722	468284	6010907
Klaipėda	5	752	323400	6174934
Šiauliai	11	396	456608	6200104
Panevėžys	9	527	522840	6177229
Kaunas	3	641	499666	6085879
Laukuva	6	642	388910	6166667
Vilnius	17	783	571223	6055189

Galima paklausti, koks geriausias būdas vizualizuoti erdvinį duomenų kitimą? Vienas atsakymas būtų: „sudaryti geografinį žemėlapi“. Šiuo atveju žemėlapyje gali būti naudojami tarpiniai taškai be stebėjimo duomenų, tačiau su prognozuojamomis reikšmėmis. Šios vietovės, su duomenimis ir be jų, gali būti vizualizuotos tinklelyje, apimančiame nagrinėjamą sritį (4.1 pav.).



4.1 pav. Geokoduotų Lietuvos meteorologinių stočių tinklas pagal lentelės duomenis ir interpoliuotas paviršius

Prognozuojant reikšmes tarpinėse vietose, gali būti naudojami interpoliacijos metodai, pvz., kriginas. Norint gauti tikslius rezultatus, gali būti taikoma optimali erdvinė prognozė ir atitinkami principai. Pvz., svertiniai pasiekiamų stebėjimų vidurkiai gali būti naudojami su didesniais svoriais, atitinkančiais artimumą prognozės taškui (4.2 pav.).



4.2 pav. Stebėjimai ir nežinomos interpoliacijos vietos

4.1.2 Kas yra geostatistika?

Geostatistika yra taikomosios statistikos šaka. Geostatistikos principus sukūrė G. Matheron (1963, Prancūzija), L. S. Gandin (1963, SSRS) ir A. S. Goldberger (1962, JAV). Iš pradžių geostatistika buvo naudojama rūdos kokybės pokyčiams kasykloje įvertinti. Vėliau principai buvo pritaikyti daugelyje geologijos sričių ir kitose mokslo srityse, įskaitant orų prognozavimą.

Unikalus geostatistikos aspektas yra **regionų kintamųjų** (*regionalized variables*), kurie patenka tarp atsitiktinių kintamųjų ir visiškai determinuotųjų kintamųjų, panaudojimas. Regionų kintamieji aprašo darinius, naudodami *geografinį skirstinį* (pvz., žemės paviršiaus aukštį ar temperatūrą). Regiono duomenyse užfiksuotos ir saugomos vietos.

Nors stebimas darinys gali būti tolydus erdvėje, ne visada įmanoma gauti duomenis iš visų vietų. Todėl nežinomos reikšmės turi būti įvertintos, atsižvelgiant į tiriamos vietos dydį, formą, orientaciją ir erdvinę tvarką. Jei pasikeičia erdvinė struktūra, pasikeičia ir įvertintos reikšmės. **Geostatistiniams** duomenims pavyzdžių atranka ir *regionų kintamųjų* įvertinimų sukuria struktūrą, kuri žemėlapyje gali būti pateikta kaip rastrinė geografinė matrica ar horizontalės.

Todėl geostatistika analizuoja geografiškai orientuotą darinį, gautą atskirais matavimais. Analizuojamo darinio tolydžiam modeliui formuoti, naudojant interpoliacijos, glodinimo ir (arba) prognozavimo technikas, geostatistika naudoja erdvines koordinatas. Nurodytos technikos naudoja erdvinių koordinatų ir empirinių duomenų pasiskirstymo informaciją.

Geostatistiniai duomenys gali padėti atsakyti į tokius klausimus:

- Kur? Pvz., kur Lietuvoje yra vandens kokybės matavimo stotys?
- Kiek? Pvz., koks nedarbo lygis Lietuvos savivaldybėse?
- Kiek? Pvz., kokios dirvožemio pH reikšmės Lietuvoje, konkrečiose vietovėse?

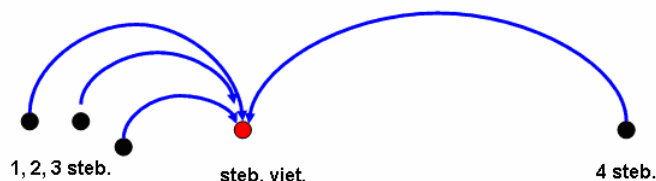
4.1.3 Pirmasis geografinės dėsnis

Kuo erdviniai duomenys ypatingi?

- Bandinio **vieta** yra esminė jo apibrėžimo dalis. Erdviniai duomenys dažnai reprezentuoja **vieno** atsitiktinio kintamojo stebėjimus. Pvz., mirtingumo koeficiento, dirvožemio pH ar temperatūros Lietuvoje stebėjimus.
- *Koreliacijos* analizėje nustatomas tiesinis sąryšis tarp **dvių** atsitiktinių kintamųjų (pvz., tarp dirvožemio pH ir javų derliaus, temperatūros ir aukščio ir pan.). Geostatistikoje naudojama pagrindinė **autokoreliacijos** funkcija arba koreliacija tarp stebėjimų, atliktų išmatuotais atstumais ir kryptimis. Todėl reikšmės yra: (1) kintamojo matavimas atsitiktiniame taške ir (2) atstumo erdvėje ir krypties struktūra tarp matuojamų kintamųjų kiekviename taške. Priklausomybę tarp šių dviejų reikšmių rinkinių galima pateikti variograma (arba pusvariogramė). Šis įrankis bus aptartas vėliau šioje kurso dalyje.
- Visi duotosios srities duomenų rinkiniai yra visiškai susiję koordinatėmis ir naudojami modeliuoti erdvinei struktūrai. Kadangi gali egzistuoti duomenų erdvinė struktūra, reikšmės bandinio taškuose *negali* būti laikomos **nepriklausomomis**. Tai skirtumas nuo *klasikinės statistikos*, kurioje bent jau bandinio duomenų reikšmės laikomos nepriklausomomis. Šios skirtingos prielaidos (klasikinės statistikos ir geostatistikos) turi didelės įtakos atrankos schemai ir statistiniams apskaičiavimams.

Erdvinius duomenis charakterizuoja priklausomybė arba autokoreliacija erdvėje. **Erdvinės priklausomybės** gali būti aprašytos „*pirmuoju geografinės dėsnio*“, kuris teigia, kad „*viskas yra tarpusavyje susiję, bet arčiau esantys dalykai yra labiau susiję nei tolimesni*“ (Tobler, 1970). Todėl, didėjant atstumui, dviejų vietų charakteristikos tampa mažiau panašios ir (arba) susijusios.

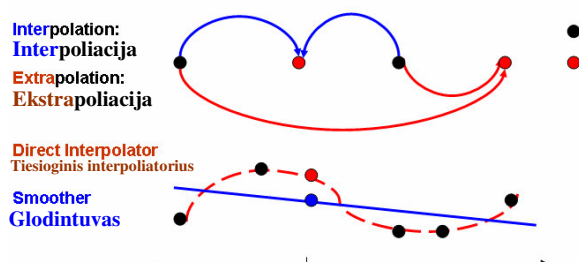
Pagrindinis geostatistikos klausimas – „*kaip galima matuoti erdvinę priklausomybę?*“ Svarbūs matavimo įrankiai yra pusvariogramė ir autokovariacijos funkcijos. Šių įrankių panaudojimas bus aptartas vėliau. Variograma yra efektyvus įrankis prognozuoti neišmatuotų taškų reikšmėms, su išmatuotais taškais naudojant atstumą ir kryptį, ir atrinkti tokios informacijos klasterius.



4.3 pav. Geostatistiniais metodais įvertinami kintamieji yra esamų stebėjimų reikšmių svertiniai vidurkiai. Stebėjimui, esančiam arčiau (t. y. turinčiam didesnę koreliaciją, nebūtinai mažiau nutolusiam) ieškomo taško, turi būti suteiktas didesnis svoris. Atrankos efektas: stebėjimų klasterių svoris sumažinamas.

4.1.4 Užduotys: interpoliacija, glodinimas ir prognozavimas

Šioje kurso dalyje aptariamos keturios technikos, geostatistikoje naudojamos kurti tolydiems paviršiams iš atrankos taškų, išdėstomų vietovėje. Šios technikos dažnai vadinamos *interpoliacija*, tačiau, kalbant griežtai, tai galioja tik taškams, geografiškai patenkančioms į ribas, apibrėžiamas atrankos taškų (priešingu atveju tai yra *ekstrapoliacija*). Kalbant apie geostatistiką, šios technikos gali būti laikomos interpoliatoriumi, glodintuvu, įvertinimo arba prognozės įrankiu, atsižvelgiant į konkrečią matematinę ir (arba) statistinę techniką. Interpoliacijos metodai gali būti panaudoti ekstrapoliacijai (4.4 pav.).



4.4 pav. Glodinimas ir tiesioginė interpoliacija

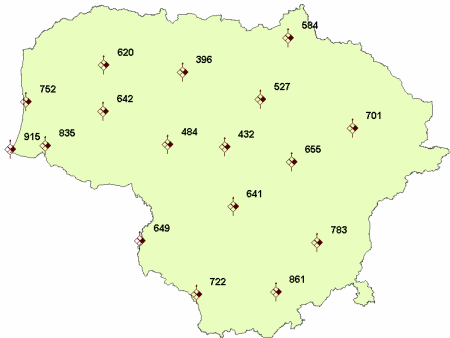
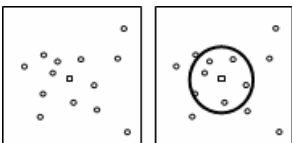
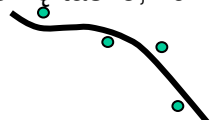
1. **Atvirkščiai proporcingo atstumo** (IDW) metodas yra determinuotasis, tiesioginis interpoliatorius.
2. **Trendo paviršiaus įvertinimas** yra stochastinis *glodintuvas*, pagrįstas globaliais ar lokaliais polinomais. Trendo paviršiaus analizė vertina tiesinį trendą, t.y. $\mu(s_i)$ gali būti determinuotasis stochastinio proceso komponentas.
3. **Splainai (dalimis tiesiniai polinomiali)** yra determinuotasis *glodintuvas*, veikiantis su lokaliais ar globaliais kaimynais.
4. **Krigingas** yra stochastinis *prognozės įrankis*, galintis veikti kaip *glodintuvas* ar tiesioginis interpoliatorius ir priklausantis nuo variogramos ar duomenų. Duomenys $z(s_i)$ yra atsitiktinių kintamųjų $Z(s_i)$ stebėjimai s_i vietoje nagrinėjamoje srityje D . Suformuojamas stochastinis procesas $\{Z(s_i) : s_i \in D, i = 1, 2, \dots\}$. Šis procesas susideda mažiausiai iš dviejų komponentų $Z(s_i) = \mu(s_i) + \eta(s_i)$ „trends ir liekana“ arba trijų komponentų $Z(s_i) = [\mu(s_i) + \eta(s_i)] + \varepsilon(s_i)$ „signalas ir triukšmas“, kur

$\mu(s_i)$ = determinuotasis trendas,
 $\eta(s_i)$ = stochastinė, koreliuojanti liekana,
 $\varepsilon(s_i)$ = atsitiktinis triukšmas, nekoreliuojantis.

4.2 Geostatistinių metodų klasifikacija

Geostatistinės technikos ar metodai gali būti suskirstyti į kategorijas pagal tris kriterijus, pateiktus 4.2 lentelėje.

4.2 lentelė. Geostatistinių technikų skirstymas į kategorijas

Globalus metodas arba Lokalus metodas	Globalus metodas naudoja visus žinomus ar bandinio taškus įvertinti  nežinomai reikšmei Lokalus metodas naudoja taškų su žinomomis reikšmėmis imtį įvertinti nežinomai reikšmei. Apskaičiavimams randamas ir naudojamas  arčiausiai esantis N taškas.
Tikslus arba Netikslus	Ne. Tiksli interpoliacija nustato reikšmę taško vietoje tokią, kaip žinoma reikšmė. Kitaip sakant, tiksli interpoliacija generuoja paviršių, einantį per kontrolinius taškus.  Taip. Netiksli interpoliacija arba apytikslė interpoliacija prognozuoja reikšmę taške, kuri skiriasi nuo žinomos reikšmės.  Ar skiriasi išmatuota ir interpoliuota reikšmės imties taškų vietose?
Determinuotasis arba Tikimybinis	Determinuotieji interpoliatoriai atlieka prognozes, naudodami matematinės formules, formuojančias gretimų žinomų reikšmių svertinius vidurkius. Determinuotasis interpoliacijos metodas neatlieka prognozuojamų reikšmių paklaidų įvertinimo. Skirtingi metodai naudoja skirtingus būdus formuoti svertiniams vidurkiams. Ši grupė apima atvirkščiai proporcingo atstumo metodą, globalius ir lokalius polinomus bei spindulio tipo funkcijas ar splineus. Stochastinės interpoliacijos taip pat naudoja svertinius vidurkius, tačiau prognozėms atlikti naudojami ir tikimybiniai modeliai. Stochastiniai interpoliacijos metodai siūlo prognozės paklaidų įvertinimą su numatytais skirtumais. Ši grupė apima krigingą ir visus antrinius jo metodus.

Šioje kurso dalyje aptariami metodai gali būti suskirstyti taip:

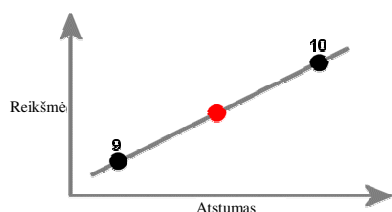
Globalūs		Lokalūs	
Determinuotieji	Stochastiniai	Determinuotieji	Stochastiniai
Trendo paviršius (netikslus)	Regresija (netikslus)	Atvirkščiai proporcingo atstumo metodas (tikslus) Splainai (tikslus)	Krigingas (tikslus)

4.3 Interpoliacija, vertinimas ir glodinimo metodai

4.3.1 Atvirkščiai proporcingo atstumo (IDW) interpoliacija

Kaip nurodėme pirmiau, interpoliacija yra nežinomų reikšmių, patenkančių tarp žinomų, įvertinimo procesas. Taškai su žinomomis reikšmėmis vadinami žinomais taškais, kontroliniais taškais, atrankos taškais ar stebėjimais. Reikšmės gali apibūdinti bet kokį kiekybinį geografinį reiškinį. Erdvinės interpoliacijos tikslas – sukurti paviršių, modeliuojantį šį reiškinį pačiu geriausiu būdu.

Interpoliacija veikia tik tada, kai reikšmės yra priklausomos erdvėje arba autokoreliuoja erdvėje, t.y. gretimoms vietoms turi panašias Z reikšmes. Autokoreliuojančių erdvėje elementų pavyzdžiai yra aukštis, nuosavybės reikšmė, nusikalstamumo lygis ir krituliai. Ne-autokoreliuojančių elementų pavyzdys yra cepelinai, suvalgomi viename namų ūkyje. Kai kraštovaizdžio reikšmės yra geografiškai nepriklausomos, interpoliacija neveikia, nes taško, esančio vietoje (x, y) reikšmė negali būti naudojama nustatyti taško, esančio $(x+1, y+1)$ reikšmei.



4.5 pav. Vienmatė ir dvimatė interpoliacija: Nežinoma reikšmė yra interpoliuojama pagal reikšmes ir atstumus iki gretimų kontrolinių taškų

Atvirkščiai proporcingo atstumo metodas naudoja pirmojo geografinės dėsnio principą. IDW apskaičiuoja gardelių reikšmes, naudodamas tiesinę taškų svorių kombinaciją. Ši technika įvertina Z reikšmę nežinomame taške, artimesniems taškams priskirdama didesnę svorį. Taip sukuriamas atvirkštinis ryšys tarp svorio ir atstumo. Matematiškai tai gali būti aprašyta toliau pateikta Shepard IDW formule (Shepard, 1968):

$$z_j = \frac{\sum_{i=1}^n \frac{z_i}{d_{ij}^p}}{\sum_{i=1}^n \frac{1}{d_{ij}^p}}, \text{ kur:}$$

z_j yra įvertinta reikšmė taške (x_j, y_j) ,

z_i yra kaimyninė reikšmė taške (x_i, y_i) ,

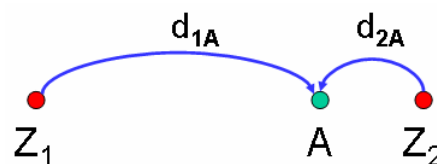
d_{ij} yra atstumas tarp (x_i, y_i) ir (x_j, y_j) ,

p yra laipsnis,

$i \in [1, n]$, n yra duomenų taškų skaičius z_j kaimynystėje.

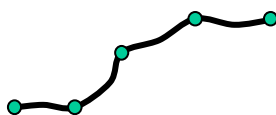
IDW tiesinės interpoliacijos atveju (laipsnis $p = 1$) žinomos reikšmės taškuose z_1 ir z_2 , atitinkantys atstumai iki taško A yra d_{1A} ir d_{2A} . Nežinoma reikšmė A taške bus apskaičiuota:

$$z_A = z_1 + \frac{d_{1A}}{(d_{1A} + d_{2A})}(z_2 - z_1) = \frac{z_1 \times w_1 + z_2 \times w_2}{w_1 + w_2}$$



$$\text{Svoriai yra } w_1 = \frac{1}{d_{1A}} \text{ ir } w_2 = \frac{1}{d_{2A}}, \sum_{i=1}^n w_i = 1$$

IDW savybės: IDW gali suteikti gerą preliminarų interpoliuojamo paviršiaus aprašą. Duomenims prielaidų nereikia, tačiau nėra prognozavimo paklaidų įvertinimo. IDW geriausiai veikia kai atrankos taškai yra tankūs, tolygiai išdėstyti. Jis neatpažįsta trendo duomenyse ir negali nustatyti reikšmių, didesnių ar mažesnių nei atrankos taškų reikšmės (4.6 pav.).

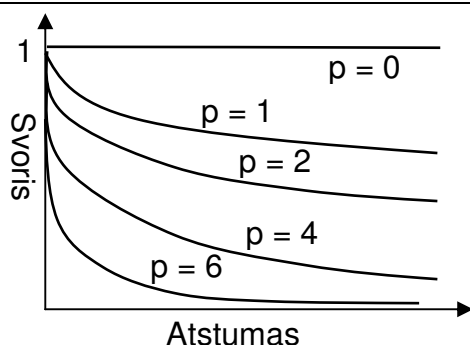


4.6 pav. IDW gautas paviršius eina per atrankos taškus, didžiausia ir mažiausia interpoliuojamo paviršiaus reikšmės gali būti tik atrankos taškuose

IDW modelio parametrai: Naudojant IDW interpoliaciją interpoliuoto paviršiaus charakteristikos gali būti valdomos nurodytais parametrais:

- **Laipsninės funkcijos** reikšmė
- **Paieškos kaimynystėje strategija**, ribojančia įvedamų taškų, kurie gali būti naudojami apskaičiuojant visų interpoliuojamų taškų reikšmes, skaičių. Kontrolinių taškų riba gali būti apibrėžta:
 - Arčiausiai esančių naudojamų atrankos taškų **skaičiumi**
 - **Paieškos spinduliu**, naudojamu paieškai kaimynystėje
 - Paieškos kaimynystėje **forma**
 - Minėtų strategijų deriniu

Laipsninė funkcija: Atsižvelgiant į vietovės sąlygas, atstumo svoris gali būti nustatytas keliais būdais. Jei, $p = 1$, tai paprasta tiesinė interpoliacija tarp taškų. Daugeliu atveju pastebėta, kad, kai $p = 2$, gaunami geresni rezultatai. Šiuo atveju artimi taškai turi didelį svorį, o labiau nutolę taškai turi nedidelį svorį (taškų svoris – $1/d_{ij}^2$). Kitais atvejais buvo naudojami kiti p laipsniai, tokiu būdu buvo gauti neblogi rezultatai. Keičiant laipsnį, galima reguliuoti santykinę atrankos taškų įtaką. *Padidintas* laipsnis reiškia, kad išvedamos reikšmės yra labiau lokalias, o ne apskaičiuotos kaip visų kaimynystėje esančių atrankos taškų reikšmių vidurkis. *Mažinant* laipsnį, išvedamos artimesnės vidurkiui reikšmės, nes toliau esantys atrankos taškai turi vis didesnę poveikį, kol visų atrankos taškų įtaka tampa vienoda.

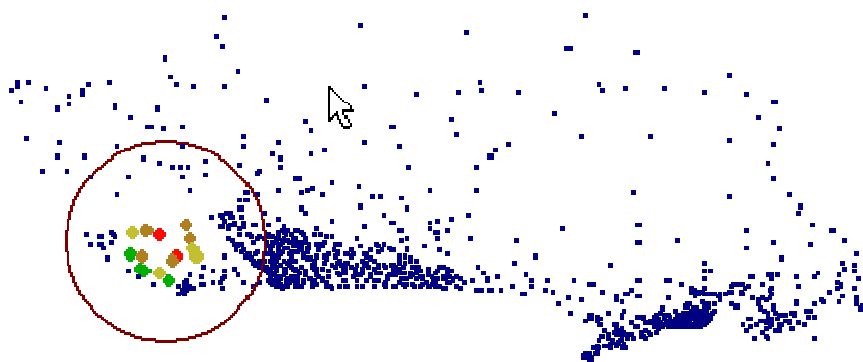


4.7 pav. Atstumo ir svorio funkcijos

Paieškos kaimynystėje strategija: Įprasta apibrėžti paieškos ribas, norint apriboti išmatuotų reikšmių, kurios naudojamos apskaičiuojant visų interpoliuojamų taškų reikšmes, skaičių. Nuo artimiausių kaimynų taškų skaičiaus priklauso imties dydis. Tokiu atveju tik N artimiausių kontrolinių taškų bus naudojami interpoliacijai nežinomoje vietoje.

Paieškos spindulys kartu su kaimyninės srities forma, apibrėžiančia paieškos ribas, nustato imties dydį. Apskaičiuoti nežinomai taško reikšmei gali būti naudojami kai kurie arba visi atrankos taškai nurodytu spinduliu.

Naudodami kitus parametrus, galime tikslinti vietų kaimynystėje pasirinkimo apribojimus. Gali būti naudojamas fiksuotas paieškos spindulys. Tuomet bus naudojami visi šiuo spinduliu esantys taškai, neatsižvelgiant į skaičių. Tokiu atveju ieškoti taškams aplink visas interpoliuojamas vietas naudojamas apskritimo *spindulio ilgis* ir *mažiausias ar didžiausias taškų skaičius*, kiek jų reikia surasti (4.8 pav.).

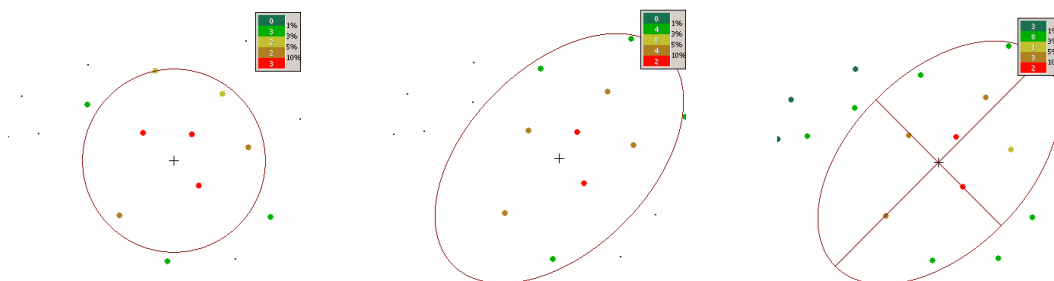


4.8 pav. Paieškos spindulio strategija – interpoliacijai naudojama ne daugiau kaip 15 artimiausių taškų, esančių paieškos spindulio atstumu

Paieškos spindulys gali būti *kintamas*. Kintamas paieškos spindulys plėsis, kol bus sudarytas nurodytas imties dydis. Galima nurodyti *skaičių taškų*, kurių reikia ieškoti skaičiuojant interpoliuojamo elemento reikšmę ir paieškos spindulio *didžiausią ilgį*.

Be to, galima keisti paieškos spindulio *formą* ir *struktūrą*. Priklausomybės erdvėje (autokoreliacija) gali priklausyti tik nuo atstumo tarp dviejų vietų, tai vadinama **izotropija**. Jei kryptinio poveikio duomenims nėra (*izotropija*), tuomet kaimynystė gali būti apibrėžta *apskritimu*.

Tačiau įmanoma, kad ta pati autokoreliacijos reikšmė gali būti nustatoma skirtingais atstumais skirtingomis kryptimis. Šiuo atveju pokyčiai vienomis kryptimis didesni nei kitomis. Ši kryptinė įtaka, vadinama **anizotropija**, galima pamatyti semivariogramoje. Jei yra kryptinė įtaka duomenims (*anizotropija*), kaimynystė gali būti apibrėžiama *elipse* su pagrindine ašimi, einančia pokyčio kryptimi. Jei kaimynystė yra padalinta į sektorius, apribojimai gali būti taikomi kiekvienam *sektoriui* (4.9 pav.).



4.9 pav. Izotropija, anizotropija ir struktūrinė anizotropija

Svarbu išnagrinėti anizotropiją ir ją naudoti interpoliacijos modelyje kaip parametą (pvz., kryptinė paieškos kaimynystės forma). Variogramos anizotropija toliau bus aptarta detaliau.

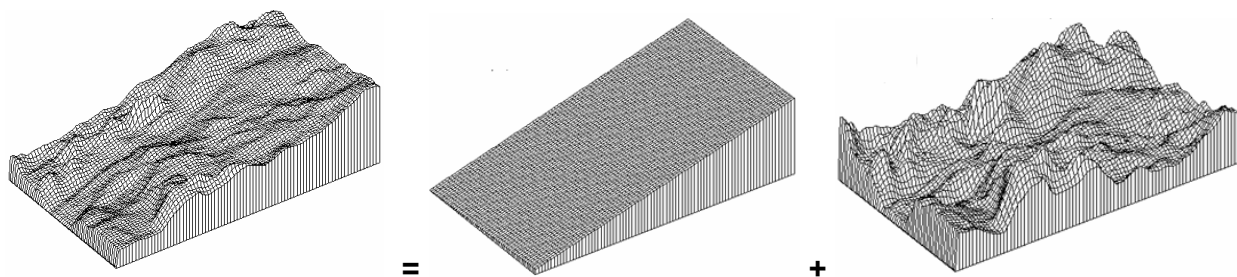
Kiti paieškos kaimynystėje apribojimai apima fizines *kliūtis*, pvz., kalnynus, dėl kurių interpoliatorius gali nenaudoti atrankos taškų kitoje to kalno pusėje. IDW metodo modifikacijos analizėje gali aptarti tokias kliūtis. Kliūtis gali būti polilinijų duomenų rinkinys, naudojamas kaip skyriklis, ribojantis įvedamų bandinio taškų paiešką.

IDW santrauka

- IDW yra dažniausiai naudojama interpoliacijos technika.
- Atvirkščiai proporcingo atstumo metodas yra:
 - geostatistinis metodas;
 - tikslus (klasikinės formos);
 - lokalus;
 - determinuotasis.
- Yra daug skirtingų Shepard IDW metodo modifikacijų, iš kurių nei viena nėra idealiai tinkama visiems taikymams. Nors skiriasi metodų svorių nustatymas (pvz., gali būti naudojamos funkcijos $\frac{1}{1+cd^2}$; e^{-cd} ; e^{-cd^2} , kur c yra konstanta) ir naudojamų stebėjimų skaičius, visos IDW modifikacijos kintamojo interpoliacijai neišmatuotose vietovėse naudoja vietovę ir reikšmę. Visais metodais, net ir turint tuos pačius duomenis, gaunami skirtingi rezultatai.
- IDW tikslumas dažnai nusakomas skirtumo tarp išmatuotų (ir (ar) kontrolinių taškų) ir interpoliuotų reikšmių vidutine kvadratine paklaida.

4.3.2 Globalūs polinomialai

Paviršiaus ypatybių analizei naudojamos globalios paviršiaus modeliavimo procedūros. Paviršius gali būti aproksimuotas kontrolinių taškų geografinių koordinatų globalių polinomų plėtimu. Polinominės funkcijos koeficientai randami mažiausiųjų kvadratų metodu, užtikrinant, kad kvadratinų nuokrypių nuo paviršiaus trendo suma yra mažiausia. Įvertintas trendas, t. y. $\mu(s_i)$ arba $Z(x,y)$ laikomas determinuotuoju stochastinio proceso komponentu.



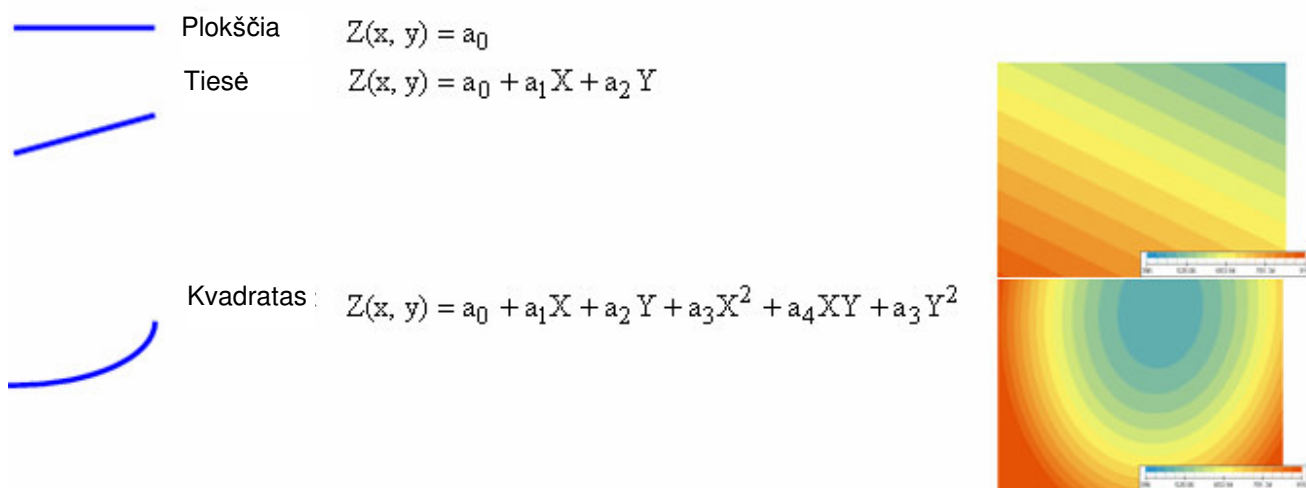
4.10 pav. Kiekvienas originalus stebėjimas laikomas determinuotosios polinominės funkcijos geografinių koordinatų bei atsitiktinės stochastinės liekanos suma.

Globali polinominė funkcija trendo analizei yra tiesinė nurodytos formos funkcija:

$$Z(x, y) = a_0 + a_1X + a_2Y + a_3X^2 + a_4XY + a_5Y^2 \dots$$

Kur $Z(x, y)$ yra duomenų reikšmė aprašytoje vietoje, a yra koeficientai, o X ir Y yra geografinės vietos kombinacijos. Polinominė trendo paviršiaus analizė iš principo yra tiesinė regresijos technika, bet ji taikoma dvimatėse ir trimatėse erdvėse, o ne linijai.

Gali būti pasirinktas bet kuris polinomo laipsnis (4.11 pav.). Praktikoje trendo analizei naudojami nedideli laipsniai (pirmos, antros ir trečios eilės). Plokščią paviršių apibūdina $Z(x, y) = a_0$. Tiesinė pirmos eilės polinominė plokštuma yra $Z(x, y) = a_0 + a_1X + a_2Y$. Antros eilės kvadratinis polinomas ir $Z(x, y) = a_0 + a_1X + a_2Y + a_3X^2 + a_4XY + a_5Y^2$ leidžia išgauti paviršių su vienu išlenkimu t.t. Nežinomi a koeficientai yra gaunami išsprendžiant tiesinių lygčių sistemą su laipsnių sumomis X , Y bei Z reikšmių vektorinėmis sandaugomis.



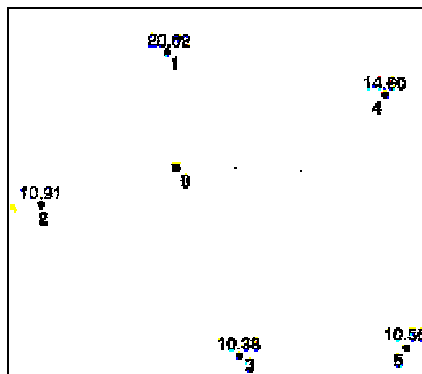
4.11 pav. Nedidelės eilės paviršiai

Optimalus trendas arba tiesinė funkcija turi kiek galima sumažinti kvadratinus nuokrypius nuo trendo. Sprendimas matricos forma:

$\mathbf{Z} = \mathbf{a} * \mathbf{B}$, kur $\mathbf{Z}$ yra tiesinė žinomų reikšmių matrica, $\mathbf{a}$ yra tiesinė žinomų $\mathbf{a}$ koeficientų matrica, $\mathbf{B}$ yra kvadratinė matrica, gauta iš geografinių vietų žinomų X ir Y reikšmių. Optimizavimo sprendimas bus $\mathbf{a} = \mathbf{B}^{-1} * \mathbf{Z}$, kur $\mathbf{B}^{-1}$ yra atvirkštinė $\mathbf{B}$ matrica. Toliau pateiktas pirmosios eilės (tiesinio) paviršiaus $Z(x, y) = a_0 + a_1X + a_2Y$ skaičiavimo sprendimo pavyzdys, naudojant *mažiausiųjų kvadratų* metodą.

Bandinio duomenys 4.12 pav.

Taškas	X	Y	Z reikšmė
1	69	76	20.820
2	59	64	10.910
3	75	52	10.380
4	86	73	14.600
5	88	53	10.560
0	69	67	?



4.12 pav. Bandinys su interpoliuojamo paviršiaus žinomomis vietomis

Šis paviršius gali būti pateikiamas kaip tiesinių lygčių sistema:

$$20.820 = a_0 + 69a_1 + 76a_2$$

$$10.910 = a_0 + 59a_1 + 64a_2$$

$$10.380 = a_0 + 75a_1 + 52a_2$$

$$14.600 = a_0 + 86a_1 + 73a_2$$

$$10.560 = a_0 + 88a_1 + 53a_2$$

Norint surasti šiuos tris $\mathbf{a}$ nežinomuosius n -tajam duomenų taškui, galima naudoti šias tris normaliąsias lygtis:

$$\sum_{i=1}^n Z_i(x, y) = n * a_0 + a_1 \sum_{i=1}^n X_i + a_2 \sum_{i=1}^n Y_i$$

$$\sum_{i=1}^n X_i Z_i(x, y) = a_0 \sum_{i=1}^n X_i + a_1 \sum_{i=1}^n X_i^2 + a_2 \sum_{i=1}^n X_i Y_i$$

$$\sum_{i=1}^n Y_i Z_i(x, y) = a_0 \sum_{i=1}^n Y_i + a_1 \sum_{i=1}^n X_i Y_i + a_2 \sum_{i=1}^n Y_i^2$$

Išsprendus šią lygčių sistemą, bus gautas geriausias dvimatės, pirmosios eilės (plokštuma) paviršiaus variantas, apibrėžtas mažiausiųjų kvadratų regresija. Tai galima užrašyti naudojant matricas:

$$\begin{bmatrix} n & \sum_{i=1}^n X_i & \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 & \sum_{i=1}^n X_i Y_i \\ \sum_{i=1}^n Y_i & \sum_{i=1}^n X_i Y_i & \sum_{i=1}^n Y_i^2 \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n Z_i \\ \sum_{i=1}^n X_i Z_i \\ \sum_{i=1}^n Y_i Z_i \end{bmatrix}$$

Su pradinėmis duomenų reikšmėmis:

$$\begin{bmatrix} 5 & 377 & 318 \\ 377 & 29.007 & 23.862 \\ 318 & 23.862 & 20.414 \end{bmatrix} * \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 67.270 \\ 5043.650 \\ 4445.800 \end{bmatrix}$$

Po pakeitimų:

$$\begin{bmatrix} 23.2102 & -0.1631 & -0.1684 \\ -0.1631 & 0.0018 & 0.0004 \\ -0.1684 & 0.0004 & 0.0021 \end{bmatrix} * \begin{bmatrix} 67.270 \\ 5043.650 \\ 4445.800 \end{bmatrix} = \begin{bmatrix} -10.094 \\ 0.020 \\ 0.3470 \end{bmatrix}$$

Tiesinė funkcija yra apibrėžta kaip $Z(x,y) = -10.094 + 0.020X + 0.347Y$. Įvertinus koeficientus, polinominė funkcija gali būti vertinama bet kuriame taške nagrinėjamoje srityje (pvz., $Z(69,67) = -10.094 + 0.020 * 69 + 0.347 * 67 = 14.535$). Nesudėtinga sukurti reikšmių tinklelio matricą, pakeičiant tinklelio koordinates polinomais ir apskaičiuojant kiekvieno mazgo paviršiaus įvertinimą. Dėl mažiausių kvadratų procedūros, jokia kita to paties laipsnio polinominė lygtis negali suteikti geresnės duomenų aproksimacijos.

Globalus polinominis paviršius palaipsniui keičiasi ir sudaro grubaus mastelio duomenų struktūras. Globali polinominė interpoliacija kuria lėtai kintantį paviršių, naudodama nedidelės eilės polinomus, kurie gali apibūdinti fizikinius procesus (pvz., taršą ir vėjo kryptį). Apskaičiuoti paviršiai yra lengvai paveikiami išskirtinių reikšmių (labai didelių arba labai mažų reikšmių, ypač kraštuose - *outliers*).

Globalus polinominis įvertinimas gali būti naudojamas paviršiaus pritaikymui prie atrankos taškų, paviršiui lėtai kintant nagrinėjamoje srityje. Be to, jis taikomas nagrinėjant ir (arba) naikinant ilgų ar globalių trendų poveikius. Trendo polinominė analizė gali padėti nustatyti globalius trendus įvesties duomenų rinkinyje. Globali polinominė aproksimacija gali būti naudojama *regresijos analizei* tarp dviejų ar daugiau kintamųjų.

Globalių polinomų santrauka

- Trendo paviršiaus analizė yra matematinis metodas, naudojamas atskirti „regioninius“ nuo „vietinių“ svyravimų.
- Globalūs polinomialai yra greitas determinuotasis interpoliatorius, netikslus glodintuvas.
- Nedaug modelio parametrų pasirinkimų (tik polinomo eilė).
- Nėra prognozavimo paklaidų vertinimo.

- Vietos, esančios duomenų kraštuose, gali labai paveikti paviršių.
- Nereikalingos prielaidos apie duomenų ypatumus.

4.3.3 Lokalūs polinomialai

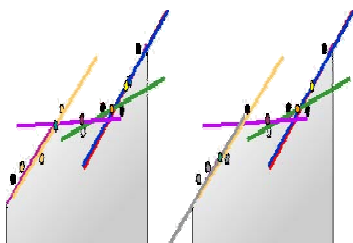
Lokalių polinomų interpoliacija yra panaši į globalių polinomų interpoliaciją, išskyrus tai, kad šiuo atveju naudojami duomenys lokalizuotose „languose“, o ne visame duomenų rinkinyje. Ji formuoja lokalius trendus „lange“ ir naudoja svorius (4.13 pav.). Langa galima judinti, o paviršiaus reikšmė $\mu_0(x_i, y_i)$, esanti lango centre, yra vertinama kiekviename taške. Turinčius svorį mažiausiuosius kvadratus galima išreikšti taip:

$$\sum_{i=1}^n w_i (Z(x_i, y_i) - \mu_0(x_i, y_i))^2,$$

$$w_i = \exp(-3d_{i0}/a),$$

Kur $\mu_0(x_i, y_i)$ yra polinomo reikšmė, d_{i0} yra atstumas tarp taško ir lango centro, o a yra parametras, kurį galima naudoti kontroliuoti svorio kitimo spartai, kintant atstumui.

Lokali polinominė interpoliacija taiko nurodytos eilės (pvz., nulinės, pirmos, antros ar trečios) polinomus, naudodami visus taškus tik apibrėžtoje kaimynystėje. Kaimynystės persidengia, o reikšmė, naudojama kiekvienai prognozei, yra pritaikyto polinomo reikšmė kaimynystės centre. Pirmos eilės polinomas yra $\mu_0(x_i, y_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i$ formos, antros eilės polinomas – $\mu_0(x_i, y_i) = \beta_0 + \beta_1 x_i + \beta_2 y_i + \beta_3 x_i^2 + \beta_4 y_i^2 + \beta_5 x_i y_i$ ir t.t. Įvyksta parametru $\{\beta_j\}$ minimizacija.



4.13 pav. Lokali polinominė interpoliacija kuria daug polinomų, visi jie su nurodytomis persidengiančiomis kaimynystėmis. Todėl lokali polinominės interpoliacijos metu gaunami paviršiai, atspindintys lokalesnius pokyčius.

Globali polinominė interpoliacija tinka kurti glodžius paviršius ir ilgų duomenų rinkinio trendų aptikimui. Tačiau žemės mokslų tyrinėjami kintamieji dažniausiai yra kintantys trumpais atstumais ir turintys tendą ilgais atstumais. Kai duomenų rinkinys pasižymi kintamumu trumpais atstumais, lokali polinominės interpoliacijos žemėlapiu gali jį pavaizduoti.

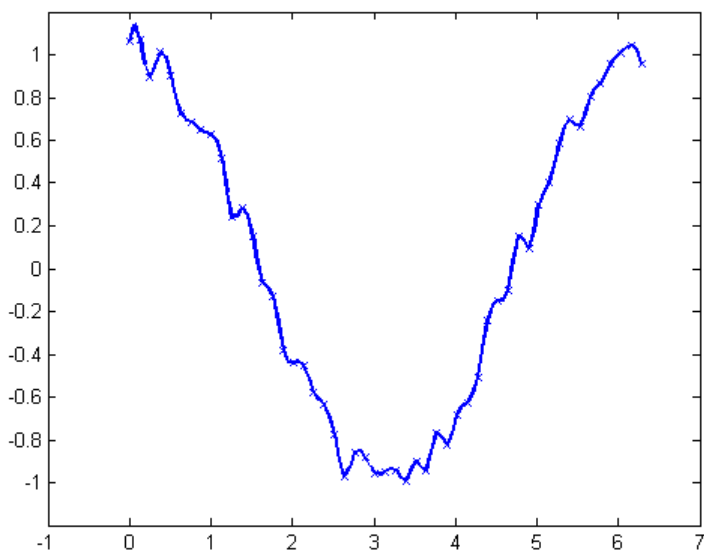
Lokali polinominė interpoliacija jautri kaimynystės atstumui. Operatorius gali interaktyviai pasirinkti šį atstumą. Kaip ir IDW atveju, galima apibrėžti anizotropiją įvertinantį modelį, parenkant tinkamą paieškos kaimynystėje formą.

Lokalių polinomų santrauka

- Lokalūs polinomialai yra greitas determinuotasis interpoliatorius, glodus ir netikslus.
- Daugiau pasirenkamų parametrų: polinomo eilė, kaimynystės dydis ir forma bei persidengimas
- Nėra prognozavimo paklaidų vertinimo.
- Metodas teikia prognozių paviršius, kuriuos galima lyginti su krigingo su matavimo paklaidomis metodu gautais paviršiais.
- Nereikalingos prielaidos apie duomenis.

4.3.4 Splainai arba spindulio tipo funkcijos (RBF)

Splainų metodai yra tikslų interpoliacijos technikų serija, pritaikanti paviršiui visas išmatuotas atrankos taškų reikšmes. Yra kelios skirtingos *splainų* funkcijos ([http://en.wikipedia.org/wiki/Spline_\(mathematics\)](http://en.wikipedia.org/wiki/Spline_(mathematics))). Visų *splainų* funkcijų pavidalas skirtingas; skirtingi ir interpoliuoti paviršiai. *Splainas* koncepcijos požiūriu panašus į guminės plėvelės užtempimą ant išmatuotų taškų reikšmių, bendrą paviršiaus kreivumą padarant minimalų (4.14 pav.). *Splainų* metodai yra dirbtinių neuronų tinklų forma.



4.14 pav. Interpoliacija naudojant *splainų* funkciją

Splainai naudojami glodiems paviršiams iš daugelio duomenų taškų apskaičiuoti. Funkcijos pateikia gerus palengva kintančių paviršių, pvz., aukščio, rezultatus. Technikos netinka, kai paviršiaus reikšmėse yra staigių pokyčių nedideliu horizontaliu atstumu.

Splainų santrauka

- Spindulio tipo funkcijos (*splainai*) yra determinuotieji tikslūs interpoliatoriai.
- Daugiau pasirenkamų parametrų nei IDW, todėl lankstesnė nei IDW.
- Nėra prognozavimo paklaidų vertinimo. Spindulio tipo funkcijos neleidžia tirti duomenų autokoreliacijos.
- Metodas teikia prognozių paviršius, kuriuos galima lyginti su tikslia krigingo forma.
- Spindulio tipo funkcijos nedaro prielaidų apie duomenis.

4.4 Interpoliavimas kriginio metodu

Matematinio *pusvariogramų* (*semivariogram*) modeliu pagrįstą metodą paviršiaus reikšmės tinklelio mazguose įvertinti pirmasis suformulavo Pietų Afrikos inžinierius D. G. Krigė. Jo vardu pavadintas kriginio metodas numato geriausius neturinčius sisteminės paklaidos tiesinius paviršiaus įverčius nurodytose vietose, remiantis prielaidomis, kad paviršius yra *stacionarus* ir kad pasirinkta tinkama pusvariogramės forma. Kriginio procedūros, įvertindamos reikšmes, aptaria paklaidų bei neapibrėžtumų dydžius. Nežinomų reikšmių skaičiavimas panašus į atvirkščiai proporcingo atstumo (IDW) ir pagrįstas žinomoms reikšmėms priskiriamais svoriais. Tačiau šie svoriai yra tik optimalūs svoriai, o svoriams skaičiuoti naudojamos pusvariogramės. Pusvariogramės svoriai priklauso nuo žinomų taškų išsidėstymo – kryptių ir atstumų.

Išmatuotos reikšmės yra tik vienas iš daugelio galimų atsitiktinio „stochastinio“ proceso rezultatų. Reikšmė Z , išmatuota kiekviename taške $\vec{s}$, yra viena iš galimų atsitiktinio kintamojo $z(\vec{s})$ reikšmių. Matuojama tik viena reikšmė – ji yra tik vienas iš galimų rezultatų proceso, kurio reikšmės gali būti įvairios. Kiekviename taške vyksta atskiri atsitiktiniai tos pačios formos procesai. Tačiau taškai gali būti susiję erdvinėje priklausomybe. Šiuo atveju taškai *nėra* nepriklausomi. Kaip visuma jie sudaro stochastinį procesą visoje srityje R (t. y. laikoma, kad tam tikram atsitiktiniam procesui, kurio rezultatas yra išmatuotos reikšmės, galioja tam tikri apribojimai, o būtent erdvinė priklausomybė). Išmatuotų reikšmių imtis $Z = \{Z(\vec{s}), \forall \vec{s} \in R\}$ vadinama regioniniu kintamuoju. Šis kintamasis begalinis dvejopai: 1) taškų skaičiumi; 2) galimų reikšmių kiekviename taške skaičiumi.

Regioninių kintamųjų teorija ryšio tarp paviršiaus taškų (autokoreliacijos) laipsniui išreikšti naudoja panašią **pusdispersės** (*semivariance*) sąvoką. Pusdispersė yra tiesiog pusė kiekvienos poros taškų, atstumas tarp kurių yra žinomas, reikšmių dispersijos. Grafikas, vaizduojantis *pusdispersės* priklausomybę nuo atstumo tarp taškų, vadinamas **variograma** (*variogram*). Šį grafiką aptarsime išsamiai.

Reikiamo paviršiaus atkūrimas pagrįstas stacionarumo principu. Stacionarumas – tai prielaida, kuri erdviųjų duomenų atveju dažnai yra pagrįsta. Stacionarumas būna dviejų tipų. Pirmasis tipas vadinamas pirmos eilės arba **vidurkio stacionarumu** (*mean stationarity*). Geostatistikoje laikoma, kad matuojamų reikšmių vidurkis vienodas visuose taškuose ir nepriklauso nuo vietos.

Antrasis tipas pusvariogramų atveju vadinamas **vidiniu stacionarumu** (*intrinsic stationarity*), o kovariacijos atveju – **antros eilės stacionarumu** (<http://en.wikipedia.org/wiki/Covariance>). Vidinis pusvariogramų stacionarumas pagrįstas prielaida, kad bet kurių dviejų taškų, esančių vienodu atstumu ir ta pačia kryptimi vienas nuo kito, reikšmių skirtumų dispersija yra vienoda. Antros eilės stacionarumas yra prielaida, kad erdvinė autokovariacija tarp bet kurių dviejų taškų, esančių vienodu atstumu ir ta pačia kryptimi vienas nuo kito, yra vienoda. Ši autokovariacija priklauso nuo atstumo tarp bet kurių dviejų reikšmių, bet nepriklauso nuo jų vietos.

Tačiau iš tikro pirmos eilės stacionarumas dažnai nėra statistiškai tikėtinas. Išmatuotų reikšmių vidurkis skirtingose srityse dažnai skiriasi arba matyti aiški trendas. Antros eilės stacionarumas taip pat dažnai abejotinas, taigi didėjant srities plotui, kovariacija dažnai neribotai didėja. Tai galima išspręsti naudojant ne pačias *reikšmes*, o jų *skirtumus*, ir apsiriboti „maža“ sritimi. Reikšmių *skirtumai* visoje srityje yra vienodi. Be to, *atėmus*

trendą, liekanos gali atitikti pirmos eilės stacionarumą. Vėliau pamatysite, kaip šie ir kiti sprendimai (pvz., empirinės variogramos vidurkių skirtumus), pritaikomi kriginio metodui.

4.4.1 Variogramos

Kiekvienai dviejuose taškuose išmatuotų reikšmių porai būdinga tam tikra dispersija. Ji lygi kintamųjų reikšmių šiuose taškuose skirtumo kvadratui. Pusdispersė – tai dispersija, padalinta iš dviejų. **Pusdispersė** apibrėžiama taip:

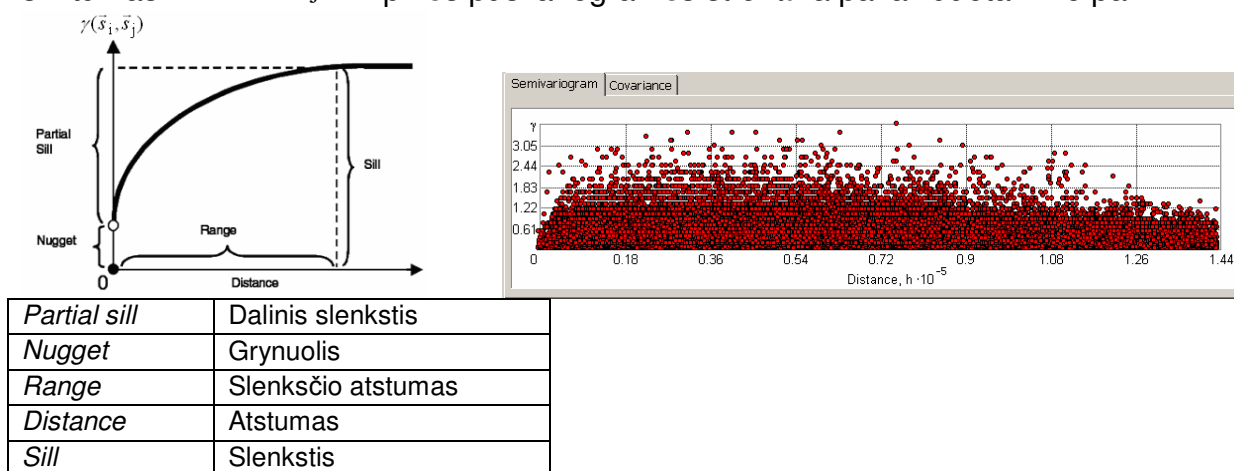
$$\gamma(\bar{s}_i, \bar{s}_j) = \frac{1}{2} [Z(\bar{s}_i) - Z(\bar{s}_j)]^2,$$

kur $\gamma(\bar{s}_i, \bar{s}_j)$ – pusdispersė, o $Z(\bar{s}_i)$ ir $Z(\bar{s}_j)$ – reikšmės taškuose $\bar{s}_i$ ir $\bar{s}_j$, atstumas tarp kurių yra žinomas. Šios formulės rezultatas – pusė kvadratu pakelto skirtumo tarp reikšmių dviejuose vietose. Pusdispersė tinka nusakyti erdvinę priklausomybę arba autokoreliaciją.

Pusdispersės galima vaizduoti **variograma**. Variograma gaunama iš turimų duomenų. Tai taškų „debesis“, grafiškai vaizduojantis dviejų to paties kintamojo reikšmių dviejuose taškuose dispersiją; jei matavimų skaičius yra n , variogramos taškų yra $n(n-1)/2$. Variogramos „debesis“ vaizduoja pusdispersės priklausomybę nuo atstumo. Abscisių (x) ašyje atidedamas atstumas tarp dviejų taškų, ordinačių (y) ašyje atidedama pusdispersė – kiekybinė autokoreliacijos išraiška.

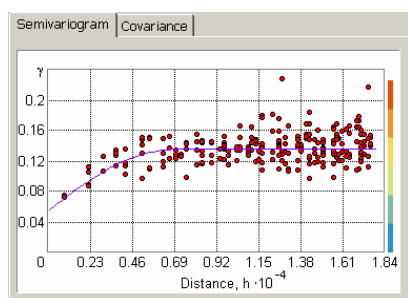
Didėjant atstumui tarp taškų, pusdispersė paprastai didėja. Variogramos dažnai apibūdinamos **grynuolio** (*nugget*), **slenksčio** (*sill*) ir **slenksčio atstumo** (*range*) parametrais. **Slenkstis** – tai didžiausia pusdispersė, t. y. didžiausias aukštis, kurį pasiekia pusvariogramės kreivė; ji atitinka kintamumą, kuris būtų, jei nebūtų erdvinės priklausomybės. Slenkstį dažnai sudaro dvi dalys: trūkis pradžioje, vadinamas **grynuolio** reiškiniu, ir **dalinis slenkstis** (*partial sill*), kuriuos sudėjus ir gaunamas **slenkstis**. **Grynuolis** yra pusdispersė, kai atstumas tarp taškų artėja prie nulio, ir atitinka reikšmių kintamumą taške, kurio nepaaiškina erdvinė struktūra. Grynuolį galima išskaidyti į matavimų paklaidą ir paklaidą dėl kintamumo mikromasteliu; kadangi kiekviena iš šių dedamųjų gali būti nulis, visą grynuolio reiškinį gali sudaryti tik viena iš jų. **Slenksčio atstumas** – tai nuotolis tarp taškų porų, kuriam esant pasiekiamas slenkstis, kitaip tariant, atstumas, kuriuo nepasireiškia erdvinė priklausomybė.

Kuo toliau taškai, tuo mažiau panašesnės reikšmės $Z(\bar{s}_i)$ ir $Z(\bar{s}_j)$ ir tuo didesnis jų skirtumas $Z(\bar{s}_i) - Z(\bar{s}_j)$. Tipinės pusvariogramės *struktūra* pavaizduota 4.15 pav.



4.15 pav. Tipinės pusvariogramės struktūra ir variogramos vaizdas ArcGIS Geostatistical Analyst programoje

Sudarant variogramą, reikšmės skaičiuojamos $n(n-1)/2$ taškų porų. Šis skaičius gali būti gana didelis ir nepatogus vaizduoti grafiškai. Pavyzdžiui, jei matuojama 200 taškų, gaunama 19 900 taškų porų. Kad **empirinės pusvariogramės** taškų skaičius būtų mažesnis, taškų poros skirstomos pagal atstumą tarp taškų į vadinamąsias **vėlinimo grupes** (*lag bin*) arba į grupes pagal tam tikrą suskaidymo atstumą $\bar{h}$. Pavyzdžiui, skaičiuojama vidutinė visų porų, kurių atstumas tarp taškų yra nuo 100 iki 200 metrų, pusdispersė. Taip atrenkamos visomis poros, kurių atstumas tarp taškų yra h , ir skaičiuojamas reikšmių skirtumo kvadratų vidurkis. Taigi, empirinė pusvariogramė yra grafikas, kurio y ašyje atidedamos suvidurkintos pusvariogramės reikšmės, o x ašyje – atskyrimo atstumas arba (vėlinimas) $\bar{h}$ (4.16 pav.).



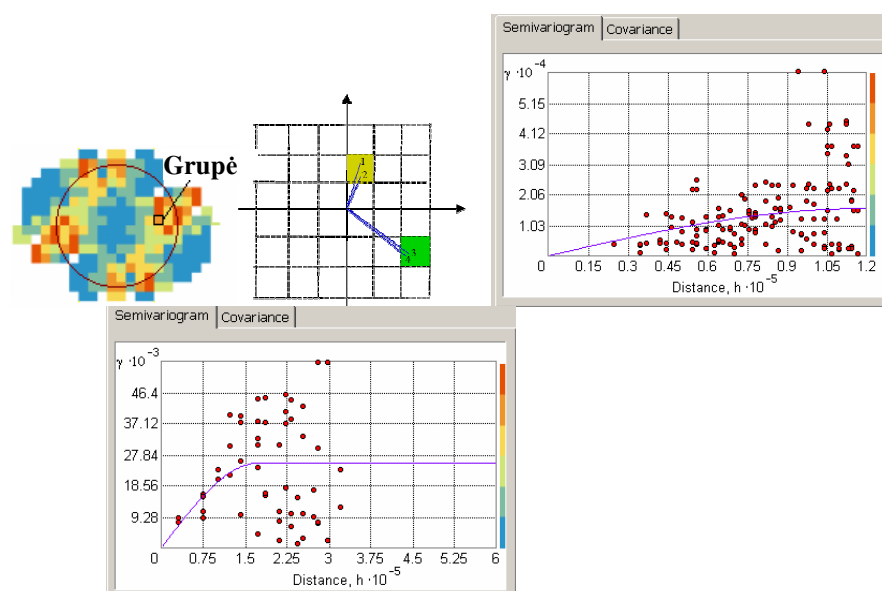
4.16 pav. Empirinės pusvariogramės vaizdas ArcGIS Geostatistical Analyst programoje. Empirinės pusvariogramės grafike kiekvienos grupės rezultatai vaizduojami vienu tašku – apskaičiuojamas visų porų vidutinis atstumas tarp taškų, kuriuo atidedama vidutinė porų pusdispersės reikšmė.

Atkreipkite dėmesį, kad skirstymas į grupes yra vidinio stacionarumo prielaida, leidžianti atkurti nagrinėjamą reiškinį. Reikšmių vidurkiai keičiami skirtumų vidurkiais, kurie visoje srityje ar bent jau esant mažam atskyrimo žingsniui $\bar{h}$. Taigi, ankstesnė pusvariogramės formulė „suvidurkinama“, ir *empirinę pusvariogramę* galima apskaičiuoti atstumams, kurie yra $\bar{h}$ kartotiniai:

$$\bar{\gamma}(\vec{h}) = \frac{1}{2m(\vec{h})} \sum_{k=1}^{m(\vec{h})} [Z(\vec{s}_i) - Z(\vec{s}_j)]^2$$
 – pusvariogramė lygi taškų porų, kurių atstumai tarp taškų patenka į $\vec{h}$ pločio grupę, reikšmių skirtumų kvadratų vidurkiui.

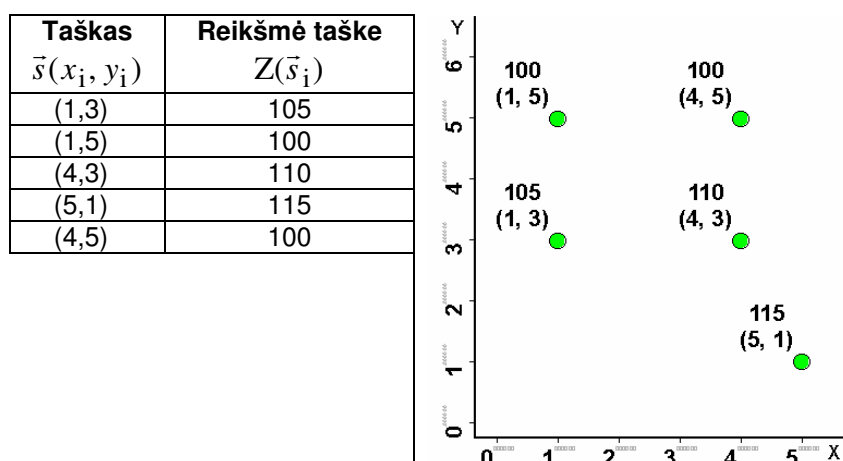
Čia $\vec{h}$ yra skirtumas tarp vienodais atstumais išdėstytų taškų (atskyrimo atstumas arba žingsnis), $Z(\vec{s}_i)$ ir $Z(\vec{s}_j)$ – reikšmės taškuose $\vec{s}_i$ ir $\vec{s}_j$, $m(\vec{h})$ – skaičius porų, kurių atstumas tarp taškų yra vektorius $\vec{h}$ arba taškų skaičius grupėje. Praktiškai, kad galėtume apibrėžti kiekvienos grupės vektorius, žinomų taškų turi būti pakankamai daug.

Grupės paprastai sudaromos skirstant tiriamąją sritį į gardeles arba sektorius, pagal kuriuos skaičiuojama empirinė pusvariogramė. Gardelės dydis vadinamas *atskyrimo žingsniu* (lag size), o gardelių skaičius – *atskyrimo žingsnių skaičiumi*. Mažinant žingsnį, didėja skiriamoji geba, bet mažėja kiekvienos imties taškų skaičius (4.17 pav.).



4.17 pav. Gauta kintamojo reikšmių panašumo (dviejų reikšmių kovariacijos) priklausomybė nuo atstumo h . Matyti (dešinėje paveikslėlio pusėje), kad pasirinkus 10 000 ir 50 000 metrų atskyrimo žingsnį (grupės dydį), gaunamos skirtingos variogramos

Empirinės pusvariogramės porų skirstymą į grupes panagrinėkime šiuo pavyzdžiu:



$$\gamma(\vec{s}_i, \vec{s}_j) = \frac{1}{2} [Z(\vec{s}_i) - Z(\vec{s}_j)]^2$$

4.3 lentelė. Pusvariogramių skaičiavimas

Taškas ($\vec{s}_i, \vec{s}_j$)	Euklido atstumo skaičiavimas $\gamma(\vec{s}_i, \vec{s}_j)$	Atstumas $d(\vec{s}_i, \vec{s}_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$	Skirtumas ² $[Z(\vec{s}_i) - Z(\vec{s}_j)]^2$	Pusdispersė $\gamma(\vec{s}_i, \vec{s}_j)$
(1,5),(4,3)	$\sqrt{(1-4)^2 + (5-3)^2}$	3,606	100	50
(1,5),(1,3)	$\sqrt{0^2 + 2^2}$	2	25	12,5
(1,5),(4,5)	$\sqrt{3^2 + 0^2}$	3	0	0
(1,5),(5,1)	$\sqrt{4^2 + 4^2}$	5,657	225	112,5
(4,3),(1,3)	$\sqrt{3^2 + 0^2}$	3	25	12,5
(4,3),(4,5)	$\sqrt{0^2 + 2^2}$	2	100	50
(4,3),(5,1)	$\sqrt{1^2 + 2^2}$	2,236	25	12,5
(1,3),(4,5)	$\sqrt{3^2 + 2^2}$	3,606	25	12,5
(1,3),(5,1)	$\sqrt{4^2 + 2^2}$	4,472	100	50
(4,5),(5,1)	$\sqrt{1^2 + 4^2}$	4,123	225	112,5

$$\bar{\gamma}(\vec{h}) = \frac{1}{2m(\vec{h})} \sum_{k=1}^{m(\vec{h})} [Z(\vec{s}_i) - Z(\vec{s}_j)]^2$$

4.4 lentelė. Empirinės pusdispersės porų skirstymas į grupes, suskirstymo žingsnis $\vec{h} = 1$ metras

Atskyrimo atstumas metrais	Porų atstumai	Vidutinis atstumas	Pusdispersė $\gamma(\vec{s}_i, \vec{s}_j)$	Vidurkis $\gamma(\vec{s}_i, \vec{s}_j)$
Nuo 1 iki 2	2, 2	2	12,5, 50	31,25
Nuo 2 iki 3	2,236, 3, 3	2,745	12,5, 12,5, 0	8,33
Nuo 3 iki 4	3,606, 3,606	3,606	50, 12,5	31,25
Nuo 4 iki 5	4,472, 4,123	4,298	50, 112,5	81,25
Daugiau negu 5	5,657	5,657	112,5	112,5

Klausimas. Kaip sužinoti atskyrimo žingsnį? Paprastai atskyrimo žingsnis pasirenkamas toks, kad jo ir žingsnių skaičiaus sandauga būtų apytiksliai lygi pusei didžiausio atstumo tarp matuojamų taškų.

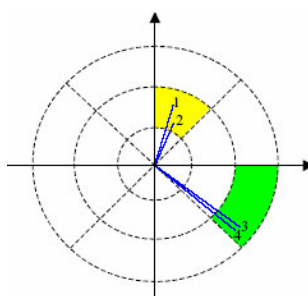
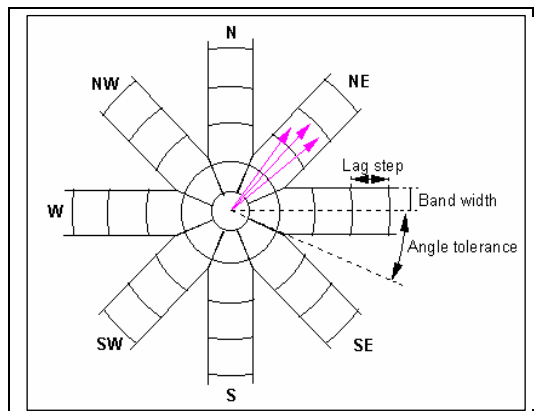
Kol kas pusvariogramės reikšmės grupuojame tik pagal *atstumą*. Tokia variograma vadinama **įvairiakrypte** ir **izotropine** (graikiškai „izo“ + „tropinė“ = lietuviškai „vienoda“ + „kryptis“). Kad supaprastintume variogramą, krypties tarp taškų, nutolusių atstumu h , kol kas nenurodėme, bet kintamumas gali priklausyti ne tik nuo atstumo, bet ir nuo *krypties*.

Paviršiaus interpoliavimo rezultatams įtakos gali turėti dvi kryptinės dedamosios:

- **globalus trendas** – stipresnis procesas, determinuotai darantis įtaką visiems matavimams. Globalų trendą galima išreikšti matematine formule (pvz., daugianariu), eliminuoti analizuojant matuojamus taškus ir pridėti vėliau, skaičiuojant galutinius interpoliavimo rezultatus. Šis procesas vadinamas trendo pašalinimu (*de-trending*);
- **anizotropija** (graikiškai „an“ + „tropic“ = lietuviškai „ne“ + „kryptis“) – kai dėl atsitiktinio proceso autokoreliacija viena kryptimi didesnė negu kita kryptimi.

Anizotropija atsiranda dėl proceso kryptingumo, pavyzdžiui, smėlio kiekio siaurose užliejamose lygumose reikšmių erdvinė priklausomybė didesnė lygiagrečia upei ašimi; antrinė mineralizacija kinta šalia intruzinio sluoksnio; gyventojų tankis skirtingas kalnuotoje vietovėje su ilgais tiesiais slėniais.

Taigi, vėl galime prisiminti *anizotropijos* reiškinį ir apibrėžti su juo susijusią **kryptinę empirinę variogramos** sąvoką. Kryptinė variograma vaizduoja erdvinį atskyrimu $\vec{h}$ atskirtų taškų erdvinį reikšmių kintamumą (4.18 pav.). Skirtumas nuo įvairiakryptės variogramos tas, kad $\vec{h}$ yra ne skaliaras, o vektorius. Krypčių skaičius gali būti įvairus. Kryptinė variograma skaičiuojama pagal tą pačią formulę, kaip ir įvairiakryptė. Atkreipkite dėmesį, kad pagal visą logiką *grynuolis* turi būti izotropinis (tai kintamumas *viename* taške).



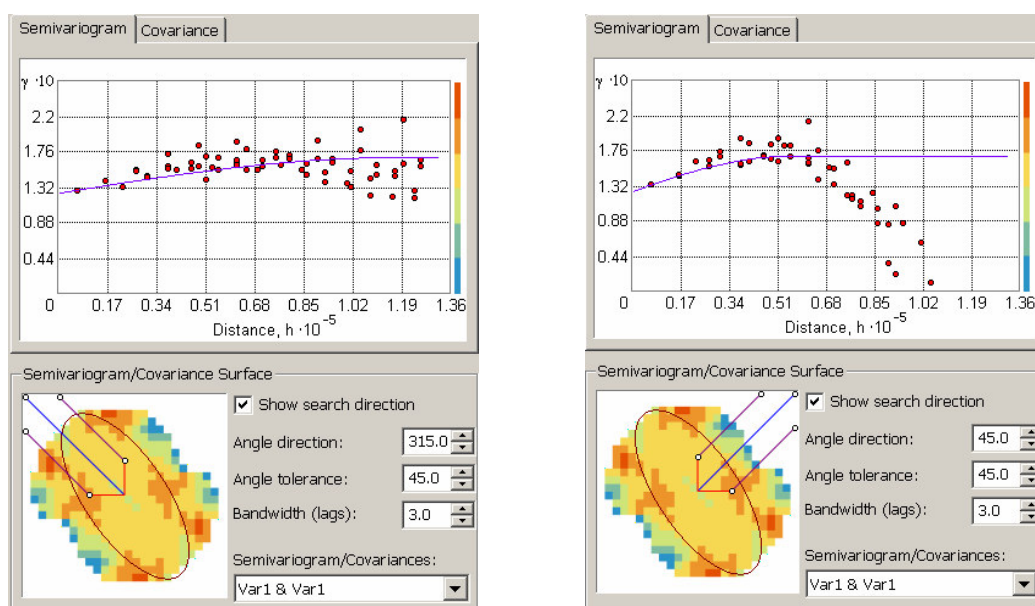
N	Š
NE	ŠR
SE	PR
S	P
SW	PV
W	V
NW	ŠV
<i>Lag step</i>	Atskyrimo žingsnis
<i>Band width</i>	Juostos plotis
<i>Angle tolerance</i>	Kampinė tolerancija

4.18 pav. Taškų poros skirstomos į grupes pagal atstumą tarp taškų ir kryptį (tam tikro pločio juostos)

Anizotropiją rodo žymūs pusvariogramių pokyčiai – gerokai skiriasi įvairiomis kryptimis apskaičiuoti slenksčio atstumai ir (arba) slenksčiai. Anizotropijos rodikliai yra dviejų tipų:

- slenksstis tas pats, bet slenksčio atstumas skirtingomis kryptimis skiriasi – tai *geometrinė* arba *afininė* anizotropija;
- slenksčio atstumas tas pats, bet slenksstis priklauso nuo krypties – tai *zoninė* anizotropija (4.19 pav.).

Daugiau informacijos apie anizotropijos tipus rasite [http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Accounting for directional influences](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Accounting%20for%20directional%20influences).

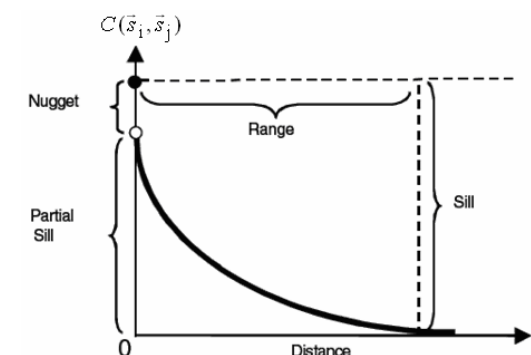


4.19 pav. Zoninė anizotropija: variograma (aprėptis) stipriai skiriasi ir priklauso nuo paieškos krypties (šiuo atveju šiaurės vakarų ir šiaurės rytų)

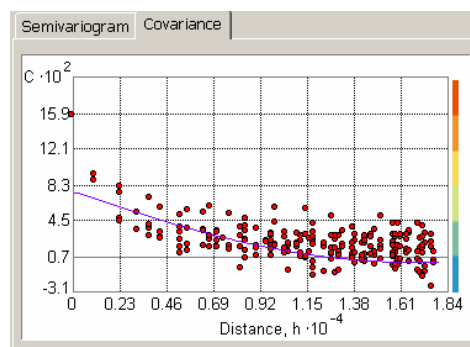
Taigi, jei nustatoma kryptinių erdvinės priklausomybės skirtumų, į juos galima atsižvelgti kuriant pusvariogramių arba kovariacijos modelius, kurie savo ruožtu turi įtakos taikomam geostatistinio interpoliavimo metodui.

Kitos erdvinės koreliacijos stiprumui kaip atstumo funkcijai įvertinti naudojamos metrikos yra **autokovariacija** ir **autokoreliacija**. Erdvinė autokovariacija skaičiuojama to paties kintamojo reikšmių poroms. Išmatuotų reikšmių poros $(\bar{s}_i, \bar{s}_j)$ autokovariacija $(Z(\bar{s}_i) - \bar{Z})(Z(\bar{s}_j) - \bar{Z})$ nurodo, kiek abi šios reikšmės kartu skiriasi nuo kintamojo vidurkio $\bar{Z}$. Erdvinę autokoreliaciją taip pat galima analizuoti kovariacinėmis funkcijomis ir korelogramomis (*correlograms*). Autokovariacijos funkcija yra $C(\bar{s}_i, \bar{s}_j) = \text{cov}(Z(\bar{s}_i), Z(\bar{s}_j))$, kur *cov* yra kovariacija. Daugiau informacijos apie sklaidos metrikas rasite <http://www.spatialanalysisonline.com/output/html/Measuresofspread.html>.

Kovariacija priklauso nuo atstumo tarp taškų. Kovariaciją reikia išreikšti kaip atskyrimo atstumo erdvėje *autokovariacijos funkciją*. Tai padarius, galima numatyti bet kurių dviejų taškų kovariaciją.



<i>Partial sill</i>	Dalinis slenkstis
<i>Nugget</i>	Grynuolis
<i>Range</i>	Slenksčio atstumas
<i>Distance</i>	Atstumas
<i>Sill</i>	Slenkstis



20 pav. Autokovariacijos funkcijos struktūra ir autokovariacijos funkcijos vaizdas ArcGIS Geostatistical Analyst programoje

Vieno kintamojo koreliacija vadinama **autokoreliacija** (priešdėlis *auto* reiškia „savi-“, t. y. kad kintamasis yra vienas). Šiuo atveju koreliacija nustatoma pagal kitus parametrus, pavyzdžiui, koreliacija *erdvėje*, jei kintamojo reikšmės matuojamos skirtingose vietose, arba *laike*, jei matuojamas laikinė reikšmių seka. Dydis, nurodantis, kiek kintamasis *koreliuoja pats su savimi* kito veiksnio (erdvės arba laiko) atžvilgiu, vadinamas autokoreliacija.

Autokovariacija yra ta pati autokoreliacija, tik kitoje skalėje. Autokoreliacija yra ne kas kita, kaip autokovariacija padalinta iš visos dispersijos $r(\vec{s}_i, \vec{s}_j) = C(\vec{s}_i, \vec{s}_j) / \sigma^2 = C(\vec{h}) / C(0)$ ($C(0)$ dispersija, kai atstumas lygus 0); jei du taškai $\vec{s}_i$ ir $\vec{s}_j$ yra arčiau vienas kito nei tikėtina pagal jų panašumą, taigi jų autokovariacija (autokoreliacija) bus didelė. Generalinės aibės

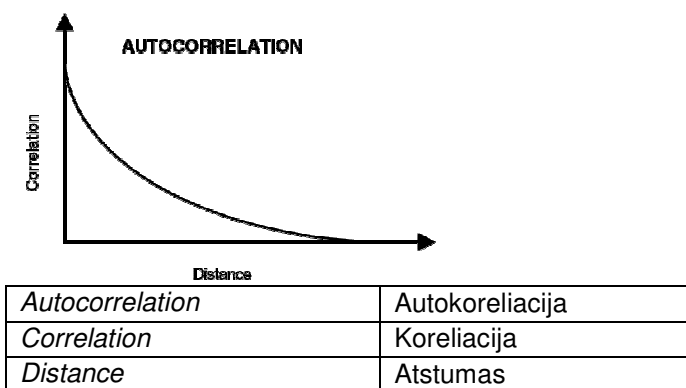
standartinis nuokrypis yra $\sigma = \sqrt{\frac{1}{n} \sum_{k=1}^n (Z(\vec{s}_k) - \bar{Z})^2}$. Didėjant atstumui tarp $\vec{s}_i$ ir $\vec{s}_j$, jie tampa mažiau panašūs ir jų kovariacija artėja prie nulio.

Ši formulė vadinama *autokoreliacijos* koeficientu atskyrimo atstumui $\vec{h}$.

$$r(\vec{h}) = \frac{\sum_{k=1}^{m(\vec{h})} (Z(\vec{s}_i) - \bar{Z})(Z(\vec{s}_j) - \bar{Z})}{\sum_{k=1}^n (Z(\vec{s}_k) - \bar{Z})^2}$$

Viršutinė šios formulės dalis panaši į kovariaciją atskyrimo atstumui $\vec{h}$, o apatinė – į kovariaciją atskyrimo atstumui 0. Šios dvi komponentės vadinamos *autokovariacija* atskyrimo atstumais $\vec{h}$ ir 0. Reikšmių rinkinį $\{r(\vec{h})\}$ galima atvaizduoti grafiškai kaip priklausomybę nuo $\vec{h}$ ir pamatyti koreliacijos kitimo keičiantis atstumui $\vec{h}$ ypatybes. Šis

grafikas (4.21 pav.) vadinamas *korelograma*. Iš jo galima gauti vertingos informacijos apie autokoreliacijos ypatybes esant įvairiems atstumams $\vec{h}$.

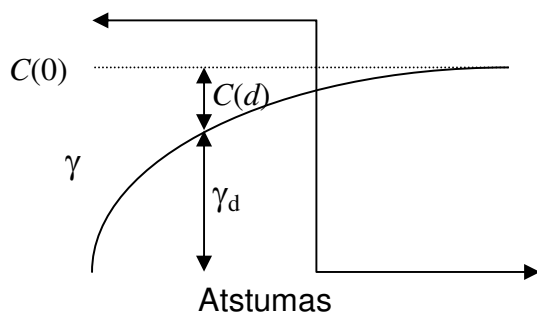


4.21 pav. Korelogramos struktūra

Reziumė:

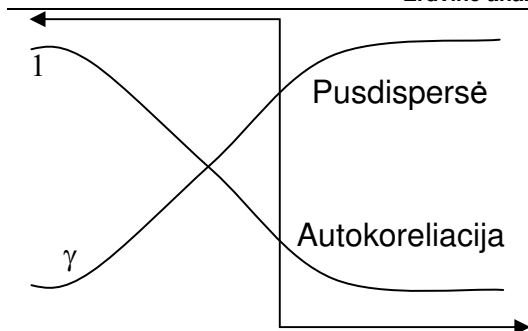
- Kovariacijos funkcija mažėja didėjant atstumui, taigi ją galima laikyti panašumo funkcija.
- Pusvariogramėje reikšmių skirtumo pusdispersė didėja didėjant atstumui, taigi ją galima laikyti nepanašumo funkcija.

Pusvariogramės susijusios su kovariacijos funkcijomis matematiniais sąryšiais, panašiais į $\gamma(\vec{s}_i, \vec{s}_j) = \text{ sill} - C(\vec{s}_i, \vec{s}_j)$ - 4.22 pav. Jei regioninis kintamasis stacionarus, pusdispersė esant atstumui d lygi dispersijos ir autokovariacijos skirtumui esant tam pačiam atstumui:



4.22 pav. Stacionaraus regioninio kintamojo pusdispersės γ ir autokovariacijos $C(d)$ ryšys. Čia $C(0)$ yra išmatuotų reikšmių dispersija, arba autokovariacija, kai $d = 0$.

Jei regioninis kintamasis ne tik stacionarus, bet ir standartizuotas taip, kad jo vidurkis yra 0, o dispersija 1, pusvariogramė yra veidrodinis autokoreliacijos funkcijos atspindys (4.23 pav.)



4.23 pav. Stacionaraus regioninio kintamojo pusdispersės ir autokoreliacijos ryšys

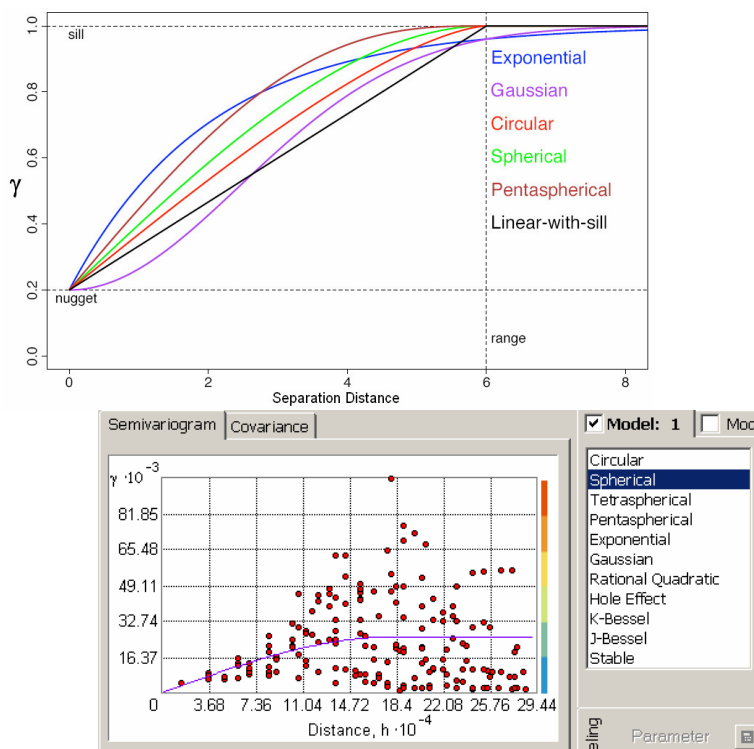
Kol kas nagrinėjome empirinės pusvariogramės ir kovariacijos „debesies“ modelius, teikiančius informaciją apie erdvinę išmatuotų duomenų rinkinių autokoreliaciją. Tačiau empiriniai pusvariogramių ir kovariacijos modeliai nesuteikia informacijos apie visas įmanomas kryptis ir visus įmanomus atstumus. Pagrindinė geostatistikos paskirtis yra atributų reikšmių numatymas (*optimalus interpoliavimas*) vietose, kuriose matavimai nebuvo atlikti.

Todėl empirinės pusvariogramės (kovariacijos) „debesiui“ reikia **pritaikyti** tolygią funkciją arba kreivę (aproksimuoti funkcija arba kreive). Ši funkcija aprašoma *teorinės variogramos* modelyje, kuriame pusdispersė išreiškta kaip *atstumo vektoriaus funkcija*. Teorinės variogramos modelis – tai funkcija, nurodanti priklausomybę tarp kintamųjų reikšmių skirtingose erdvės vietose. Laikoma, kad ši priklausomybė yra atstumo tarp reikšmių nustatymo vietų ir krypties funkcija.

Šis pritaikytas modelis naudojamas kriginio lygtyse. Abstrakčiai tai panašu į regresinę analizę, kai pritaikomos tolygios linijos arba įvairių tipų kreivės. Tačiau pusvariogramių ir kovariacijų „debesims“ aproksimuoti tinka tik tam tikros funkcijos (*leistini modeliai*). Bet kuria variogramos funkcija turi būti įmanoma modeliuoti: tolygų reikšmių didėjimą, galbūt su fluktuacija (skyle); pastovų arba asimptotinį maksimumą (slenksti), neneigiamą atkarpą (grynuolį) ir anizotropiją. Teorinės variogramos turi tenkinti tam tikrus matematinius apribojimus, užtikrinančius, kad gautos kriginio lygtys bus išsprendžiamos (pvz., teigiamai apibrėžta tarpgrupinės kovariacijos matrica).

Empirinę pusvariogramę galima modeliuoti šiomis funkcijomis: apskritinėmis, sferinėmis, tetrasferinėmis, pentasferinėmis, eksponentinėmis, Gauso, racionaliomis kvadratinėmis, skylės efekto, Beselio K ir J bei kitomis (kai kurie matematiniai modeliai aprašyti http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?topicname=how_kriging_works). Bet kuri tiesinė leistinų modelių kombinacija taip pat yra leistinas modelis.

Nuo pasirinkto modelio priklauso numatomos reikšmės, ypač jei arti koordinačių pradžios kreivių formos stipriai skiriasi. Kuo kreivė ties koordinačių pradžia statesnė, tuo didesnės įtakos numatomi reikšmei turės artimesni kaimynai. Taigi gautas paviršius bus nelygesnis.



<i>Exponential</i>	Eksponentinė
<i>Gaussian</i>	Gauso
<i>Circular</i>	Apskritiminė
<i>Spherical</i>	Sferinė
<i>Pentaspherical</i>	Pentasferinė
<i>Linear-with-sill</i>	Tiesinė su slenksčiu
<i>Sill</i>	Slenkstis
<i>Nugget</i>	Grynuolis
<i>Range</i>	Slenksčio atstumas
<i>Separation distance</i>	Atskyrimo atstumas

4.24 pav. Variogramų modelių palyginimas ir empirinės variogramos aproksimavimas sferiniu modeliu (mėlyna kreivė), gautu sukirtus atsitiktines nurodyto spindulio sferas

Tikslių taisyklių, kaip pasirinkti „geriausią“ variogramos modelį ar funkciją, nėra. Kiekvienas modelis kuriamas toks, kad kuo tiksliau atitiktų konkretaus tipo reiškinių. Pavyzdžiui, Gauso modelį galima rinktis reiškiniui, kuris fiziškai turi būti labai tolydus, pavyzdžiui, požeminio vandens gylis). Tinkamai atrodantį modelį galima pasirinkti remiantis eksperto atliktu pusvariogramės arba kovariacijos funkcijos įvertinimu, taip pat galima vadovautis patikros ir kryžminės patikros statistika (aptarsime šioje kurso dalyje vėliau).

Panagrinėkime teorinės variogramos pavyzdį. Teorinės variogramos reikia todėl, kad empirinės pusvariogramės neįmanoma tiesiogiai panaudoti matricoms skaičiuoti kriginio sistemoje (aptarsime kitame skyriuje). Naudojant empirinių variogramų reikšmes, galima gauti neigiamas numatytų reikšmių standartinės paklaidas. Todėl pusvariogramių reikšmes įvairiuose atstumuose reikia nustatyti pagal leistinus pritaikytus modelius. Šie leistini modeliai sukurti taip, kad numatytų reikšmių standartinės paklaidos niekada nebūtų neigiamos.

Ankstesniame pavyzdyje *empirinė* variograma buvo skaičiuojama taip:

Atskyrimo atstumas metrais $(\vec{h})$	Empirinė pusdispersė $\gamma(\vec{h})$
Nuo 1 iki 2	31,25
Nuo 2 iki 3	8,33
Nuo 3 iki 4	31,25
Nuo 4 iki 5	81,25
Daugiau negu 5	112,5

Pavyzdžiui, paprastas tiesinis leistinas modelis gali būti $\gamma(\vec{h}) = c * \vec{h}$, kur c yra konstanta, nurodanti teorinės variogramos tiesės nuolydį. Pritaikius tiesę suvidurkintų $\gamma(\vec{s}_i, \vec{s}_j)$ reikšmių iš lentelės regresinės analizės būdu, gaunama c reikšmė yra 13.5. Taigi teorinis modelis yra $\gamma(\vec{h}) = 13.5\vec{h}$.

Pagal šį modelį skaičiuojamos teorinės pusdispersės tarp žinomų taškų $\gamma(\vec{s}_i, \vec{s}_j)$ ir surašomos į lentelę:

Taškas $(\vec{s}_i, \vec{s}_j)$	Atstumas $d(\vec{s}_i, \vec{s}_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$	Teorinė pusdispersė $\gamma(\vec{s}_i, \vec{s}_j) = 13.5\vec{h}$
(1,5),(4,3)	3,606	48,681
(1,5),(1,3)	2	27
(1,5),(4,5)	3	40,5
(1,5),(5,1)	5,657	76,3695
(4,3),(1,3)	3	40,5
(4,3),(4,5)	2	27
(4,3),(5,1)	2,236	30,186
(1,3),(4,5)	3,606	48,681
(1,3),(5,1)	4,472	60,372
(4,5),(5,1)	4,123	55,6605

Koncepcinis variogramų metodo apibendrinimas:

- Kiekybinis duomenų reikšmių pasiskirstymo matas yra statistinis parametras, vadinamas dispersija. Dispersija yra reikšmės neapibrėžtumo matas. Krigingo metode ir kiekviena žinoma duomenų reikšmė, ir kiekviena trūkstama duomenų reikšmė turi savo dispersiją. Pastovios arba tiksliai žinomos reikšmės dispersija lygi nuliui. Blogiausiu atveju, kai duomenų reikšmė pasiskirsto visai negalima, normalizuotoje skalėje šios reikšmės dispersija lygi 1.
- Pagal prasmę, kaip matysime vėliau, kriginge dispersija atlieka svorio funkciją. Pavyzdžiui, turime vieną reikšmę, kurios dispersija lygi nuliui. Galima būti beveik tikriems, kad trūkstamos reikšmės šalia žinomos vietos bus gerai aproksimuojamos žinoma reikšme. Apie nežinomas reikšmes tolesniuose taškuose taip užtikrintai spręsti negalime; tolstant nuo žinomos reikšmės, neapibrėžtumas didėja.
- Kiekvienos žinomos duomenų reikšmės dispersiją galima nustatyti lygią jos neapibrėžtumui. Nežinomų reikšmių įverčiai labiau priklauso ne nuo absoliučių neapibrėžtumų verčių, o nuo santykinųjų jų pokyčių.
- Kiekviena žinoma duomenų reikšmė turi susijusią dispersijos funkciją, skirtą nustatyti duomenų reikšmių dispersijas vietose aplink žinomą reikšmę. Šios funkcijos forma yra empirinė ir gali būti keturių tipų: tiesinė, sferinė, eksponentinė arba Gauso. Šių kreivių minimumas (paprastai 0) yra

žinomame duomenų taške, o maksimumas (paprastai 1) – tam tikru atstumu (aprėptimi) nuo to taško. Laikoma, kad visoms vietoms už šios aprėpties žinoma duomenų reikšmė įtakos neturi. Taip pat kaip galima nustatyti vieną neapibrėžtumo reikšmę visoms žinomoms duomenų reikšmėms, visiems skaičiavimams galima naudoti ir vieną aprėpties reikšmę.

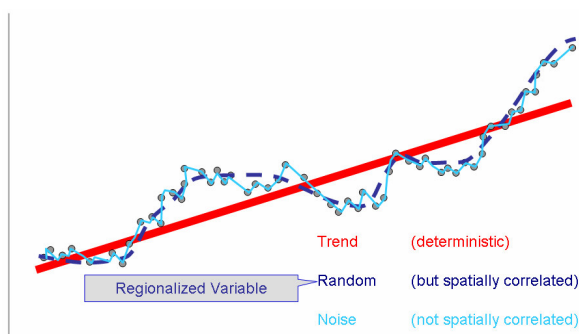
- Dėl dispersijos ir aprėpties parametrų ties žinoma duomenų reikšme gali atsirasti trūkis, dar vadinamas grynuoliu. Jis sukelia šuolinį dispersijos padidėjimą iškart už žinomos duomenų reikšmės. Geriausiu atveju ši reikšmė bus 0, t. y. grynuolio reiškinio nebus.
- Variograma *skaičiuojama* iš turimų duomenų dviem etapais:
 - pirmiausia pasirenkamos konkrečios atskyrimo žingsnio ir anizotropijos parametrų reikšmės ir skaičiuojama empirinė variograma;
 - tada pritaikomas teorinės variogramos modelis – konkreti leistina funkcija.

4.4.2 Krigingas. Regioninių kintamųjų teorija

Geostatistika unikali tuo, kad naudoja *regioninius kintamuosius* – tai kintamieji, kurie yra tarpiniai tarp atsitiktinių ir visiškai determinuotų; ši sąvoka pagrįsta prielaida, kad *regioninių kintamųjų* erdvinis kintamumas yra suma (4.25 pav.):

- struktūrinės determinuotosios dedamosios, turinčios pastovų vidurkį arba tendą; $\mu(\vec{s}_i)$
- atsitiktinės, bet koreliuotos erdvėje dedamosios ir $\eta(\vec{s}_i)$
- erdvėje nekoreliuoto atsitiktinio triukšmo. $\varepsilon(\vec{s}_i)$

Taigi, atsitiktiniai regioniniai kintamieji $z(\vec{s}_i)$ vietose $\vec{s}_i$ išreiškiami formule $z(\vec{s}_i) = \mu(\vec{s}_i) + \eta(\vec{s}_i) + \varepsilon(\vec{s}_i)$. Šiame modelyje grynuolio reiškinį gali sudaryti $\eta(\vec{s}_i)$ dispersija – vadinamasis kintamumas mikromasteliu ir $\varepsilon(\vec{s}_i)$ dispersija, dar vadinama matavimo paklaida.



<i>Trend (deterministic)</i>	Trendas (determinuota)
<i>Random (but spatially correlated)</i>	Atsitiktinė dedamoji (koreliuota erdvėje)
<i>Noise (not spatially correlated)</i>	Triukšmas (nekoreliuotas erdvėje)
<i>Regionalized variable</i>	Regioninis kintamasis

4.25 pav. Krigingo metodai priklauso nuo matematinių ir statistinių modelių Krigingo metodai pagrįsti autokoreliacijos sąvoka.

Bendra Krigingo interpoliatorių *formulė* yra svertinė duomenų suma:

$z(\bar{s}_0) = \sum_{i=1}^n \lambda_i Z(\bar{s}_i)$, kur $Z(\bar{s}_i)$ yra i-joje vietoje išmatuota reikšmė, λ_i – nežinomas i-joje vietoje išmatuotos reikšmės svoris, $z(\bar{s}_0)$ – prognozė vietoje $\bar{s}_0$, o n – išmatuotų reikšmių skaičius. Reikšmės kiekviename taške $z(\bar{s}_0)$ skaičiuojamos kaip *visuose* taškuose $Z(\bar{s}_i)$ išmatuotų reikšmių *svertiniai vidurkiai*.

Krigingo metodas panašus į atvirkščiai proporcingo atstumo (IDW) – nematuotos vietos reikšmė skaičiuojama priskiriant svorius aplinkinėms išmatuotoms reikšmėms. Abiejų interpoliatorių bendra formulė yra svertinė duomenų suma. Tačiau IDW metode svoriai priklauso tik nuo atstumų iki skaičiuojamos vietos. Krigingo svoriai gaunami iš pusvariogramės, kuri gaunama atsižvelgiant į erdvinį duomenų pobūdį – erdvinę autokoreliacija kiekybiškai išreiškiama pusdispersėmis. Taigi krigingas pagrįstas atsitiktinių procesų teorija, kai kovariacija priklauso tik nuo atstumo tarp taškų (t. y. variogramų modeliu). Nepaisant to, krigingo svoriai priklauso ne tik nuo atstumo tarp matuojamų taškų ir skaičiuojamos reikšmės vietos, bet ir nuo visų matuojamų taškų išsidėstymo erdvėje.

Be to, krigingo svoriai įvertina ne tik pusdisperses tarp interpoliuojamų ir išmatuotų taškų (IDW metode naudojami atstumai tarp žinomų ir skaičiuojamų taškų), bet ir pusdisperses tarp žinomų taškų.

Yra įvairių krigingo technikų, kurios pagrįstos įvairiomis prielaidomis. Paprastasis krigingas (*simple kriging*) yra tiesinis su žinomu trendu; normalusis krigingas (*ordinary kriging*) – tiesinis su nežinomu pastoviu trendu; universalusis krigingas (*universal kriging*) – tiesinis su nežinomu polinomine trendu; *kokrigingas* (*co-kriging*) – tiesinis ir daugiamatis, galimi įvairūs trendų tipai; transformuotas Gauso krigingas (*trans-gaussian kriging*) – tiesinis po transformacijos su pastoviu trendu; indikatorinis (*indicator*) ir disjunkcinis (*disjunctive*) krigingas – netiesinis ir tinka dirbant su trūkais arba dvejetainiais duomenimis; blokinis krigingas (*block kriging*) – tiesinis ir dirba ne su taškuose išmatuotomis, o su vidutinėmis tam tikrų mažų sričių (blokų) reikšmėmis.

Pirmiausia išnagrinėkime **normalųjį** krigingą. Normaliojo krigingo svoriai λ_i priklauso nuo išmatuotiems taškams pritaikyto modelio, atstumo iki prognozės vietos ir erdviųjų ryšių tarp aplink prognozės vietą išmatuotų reikšmių. Prognozė skaičiuojama kaip tiesinė žinomų duomenų reikšmių kombinacija (svertinis vidurkis). Normaliojo krigingo metodu gaunama prognozė yra **nepaslinktame sisteminės paklaidos**. Žinomuose taškuose apskaičiuotos reikšmės yra *tikslios*; laikoma, kad jos yra be paklaidų, net jei variogramos modelyje yra grynuolio reiškinys.

Normaliojo krigingo atveju artimesniems skaičiuojamam taškui taškams priskiriami didesni svoriai. Taškų grupės „suprastinamos“ iki vieno ekvivalentiško taško (t. y. per daug matavimų mažoje srityje rezultatų iškreipti negali). Artimesni išmatuoti taškai „užstoja“ ta pačia kryptimi esančius tolesnius. Paklaidų įverčiai priklauso tik nuo matuojamų taškų išsidėstymo ir nepriklauso nuo duomenų reikšmių.

Regioninių kintamųjų teorijos rezultatas yra „optimalus“ interpoliavimo metodas – optimalus ta prasme, kad gaunama mažiausia prognozės dispersija. Normaliojo krigingo atveju prognozės paklaida turi būti kiek įmanoma mažesnė. Kaip „geriausias nepaslinktas tiesinis prediktorius“ normalusis krigingas turi tenkinti tam tikrus optimalumo kriterijus. Tačiau „optimalus“ jis yra tik *pasirinktam pritaikytam modeliui!*

Normaliojo kriginio modeliuose reikšmė vietoje $z(\vec{s}_i) = \mu(\vec{s}_i) + \eta(\vec{s}_i)$ yra regioninio vidurkio $\mu(\vec{s}_i)$ ir erdvėje koreliuotos atsitiktinės dedamosios $\eta(\vec{s}_i)$ suma. Regioninis vidurkis $\mu(\vec{s}_i)$ apskaičiuojamas iš matavimų – bet ne kaip paprastas vidurkis, nes yra erdvinė priklausomybė. Į šią priklausomybę normaliojo kriginio sistemoje atsižvelgiama *visada*. Todėl normalusis kriginis skaičiuoja reikšmes taškuose, kai vidurkis **nežinomas** (jį reikia įvertinti) ir nėra trendo (arba jis pastovus).

Numatant kelių vietų reikšmes normaliojo kriginio metodu, tikėtina, kad kai kurios reikšmės bus didesnės už tikrąsias, o kai kurios – mažesnės. Tačiau skirtumo tarp numatytų ir tikrųjų reikšmių vidurkis turėtų būti nulis (*pirmos eilės* stacionarumo sąlyga). Tai vadinama „padaryti prognozę **be sisteminės paklaidos**“; ši sąlyga yra pagrindinis normaliojo kriginio apribojimas. Formaliai šį apribojimą galima išreikšti kaip reikalavimą, kad visiems išmatuotiems taškams priskirtų svorių λ_i suma būtų lygi vienetui. Sisteminės paklaidos nebuvimo sąlyga yra

$$\sum_{i=1}^n \lambda_i = 1$$

Dispersija kiekviename taške yra baigtinė ir vienoda visose sritys vietose; kovariacijos struktūra priklauso tik nuo atstumo tarp poros taškų (*antros eilės* stacionarumo sąlyga).

Sisteminės paklaidos nebuvimo sąlyga naudojama kartu su „optimalaus“ interpoliavimo prielaida – kad prognozės dispersija yra mažiausia įmanoma – išreiškiama taip:

$$(z(\vec{s}_0) - \sum_{i=1}^n \lambda_i Z(\vec{s}_i))^2 = \min \quad (\text{t. y. skirtumas tarp tikrosios reikšmės } z(\vec{s}_0) \text{ ir prediktoriaus}$$

$\lambda_i Z(\vec{s}_i)$ nežinomoje vietoje $\vec{s}_0$). Dispersijos minimizavimo ir sisteminės paklaidos nebuvimo apribojimo sprendinys yra normaliojo kriginio lygtys:

$$\begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_1) & \dots & \gamma(\vec{s}_1, \vec{s}_n) & 1 \\ \dots & \dots & \dots & \dots \\ \gamma(\vec{s}_n, \vec{s}_1) & \dots & \gamma(\vec{s}_n, \vec{s}_n) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \\ \Psi \end{bmatrix} = \begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_0) \\ \dots \\ \gamma(\vec{s}_n, \vec{s}_0) \\ 1 \end{bmatrix}$$

arba matricų išraiška: $A * \lambda = b$, kur $A = \begin{bmatrix} \Gamma & 1 \\ 1^T & 0 \end{bmatrix}$, $\lambda = \begin{bmatrix} \Lambda \\ \Psi \end{bmatrix}$, $b = \begin{bmatrix} \Gamma_0 \\ 1 \end{bmatrix}$. Daugiau

informacijos apie matricų išraiškas rasite <http://www.spatialanalysisonline.com/output/html/Matrixexpressions.html>.

Matricų A ir b pusdispersijų reikšmės γ_{ij} paimtos iš matematinės pusvariogramės išraiškos (*pritaikyto modelio*). Atsirado ir dar vienas kintamasis – Lagranžo daugiklis Ψ , užtikrinantis, kad gautų įverčių paklaida bus pati mažiausia. Ψ priklauso nuo išmatuotų taškų kovariacijos struktūros.

Tai $n+1$ lygčių sistema su $n+1$ nežinomųjų, taigi ją galima optimaliai suderinti, jei tik matrica A yra *teigiamai apibrėžta*, o tai **užtikrinama** naudojant *leistiną* pritaikytą modelį! Šios sistemos sprendinio matricinė išraiška:

$\lambda = A^{-1}b$, kur A^{-1} yra A atvirkštinė matrica.

Kiekvieno skaičiuojamo taško *svoris* priklauso nuo *taškų išsidėstymo* bei *variogramos modelio* ir skaičiuojamas pagal *optimizavimo* kriterijų – normaliojo kriginio atveju šis kriterijus yra *prognozės dispersijos minimizavimas*.

Terminas „normalusis“ reiškia, kad nėra trendų; regiono vidurkis apskaičiuojamas iš išmatuotų reikšmių. Vienas iš pagrindinių klausimų šiuo atveju yra prielaidos apie pastovų vidurkį teisingumas. Kartais yra pakankamai gerų mokslinių argumentų atmesti šią prielaidą.

Normaliojo kriginio metode autokoreliaciją galima išreikšti ir pusvariogramėmis, ir kovariacijomis, įvairiomis transformacijomis galima pašalinti trendus, taip pat galima atsižvelgti į *matavimų paklaidas*. Normaliojo kriginio *prognozės paklaida* arba kriginio dispersija $\hat{\sigma}$ konkrečiame taške priklauso nuo normaliojo kriginio sistemos apskaičiuotos pusdispersės tarp prognozuojamo bei išmatuotų taškų b ir nuo svorių (įskaitant Lagranžo daugiklį) λ . Normaliojo kriginio dispersija konkrečiame taške skaičiuojama pagal formulę $\hat{\sigma}^2(\vec{s}_0) = b^T \lambda$.

Dispersijos matas $\hat{\sigma}$ yra svarbus skirtumas tarp kriginio ir kitų interpoliavimo metodų (pavyzdžiui, IDW). Galima apskaičiuoti kiekvieno prognozuojamo taško $\hat{\sigma}$ ir pagal gautas reikšmes įvertinti interpoliavimo patikimumą.

Normaliojo kriginio skaičiavimus *visų* nematuotų taškų reikšmėms numatyti galima suskirstyti į šiuos etapus:

1. Apskaičiuokite atstumus tarp *kiekvienos išmatuotų taškų poros*, tada pagal *pritaikytą variogramos modelį* apskaičiuokite atitinkamas pusdisperses ir atitinkamas matricos A atstumus.

2. Apskaičiuokite atstumus nuo *prognozuojamų nežinomų taškų iki visų žinomų išmatuotų taškų*, tada pagal *teorinį variogramos modelį* apskaičiuokite atitinkamas pusdisperses ir atitinkamus vektoriaus b atstumus.

1 ir 2 kriginio sistemos etapais naudojamos sumodeliuotos (pritaikytos, teorinės) pusdispersės. Nuo pasirinkto pritaikyto modelio priklausys išmatuotiems taškams priskirti kriginio svoriai, o nuo jų priklausys prognozės rezultatai.

3. Užpildykite tuščias A ir b vietas vienetais ir nuliais.

4. Išspręskite tiesinių kriginio lygčių sistemą ir gaukite n svorių λ_i bei Lagranžo daugiklį Ψ . Šios kriginio lygtys sprendžiamos kiekvienam taškui $\vec{s}_0$ atskirai, naudojant

pusdispersės aplink tą tašką (vietinėje kaimynystėje); kiekvienam prognozuojamam taškui gausite skirtingus svorių rinkinius λ_i .

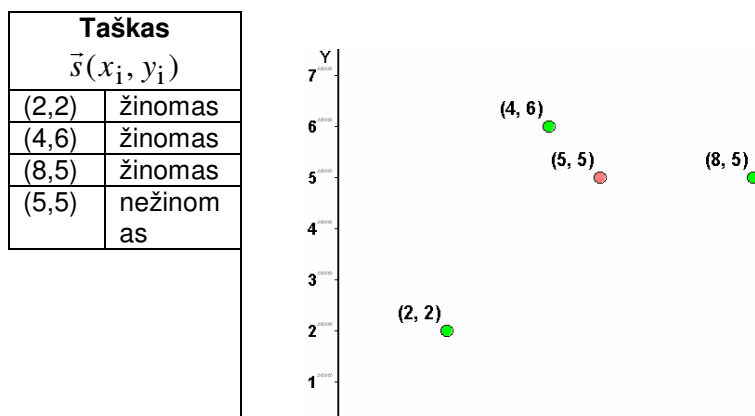
5. Pagal formulę $z(\vec{s}_0) = \sum_{i=1}^n \lambda_i Z(\vec{s}_i)$ apskaičiuokite taškų reikšmes kaip svertinius išmatuotų

taškų vidurkius su svoriais λ_i .

6. Apskaičiuokite prognozės paklaidas $\hat{\sigma}$ kaip skaliarinę $\mathbf{b}$ ir λ vektorių sandaugą.

Pastaba. Šiose lygtyse naudojamas variogramos modelis $\gamma(\vec{h})$ įvertinamas tik vieną kartą atsižvelgiant į visos tiriamos srities erdvinę struktūrą, todėl pusdispersės tarp išmatuotų taškų $\gamma(\vec{s}_i, \vec{s}_j)$ skaičiuojamos tik vieną kartą, kad ir kokia būtų taškų konfigūracija. Tačiau kiekvieno išmatuoto taško $\gamma(\vec{s}_i, \vec{s}_0)$ pusdispersė kiekvienam prognozuojamam taškui turi būti skaičiuojama atskirai.

Įsivaizduokime tokią taškų imtį:



Naudojamas eksponentinis pusdispersės modelis $\gamma(\vec{h}) = c(1 - e^{-\frac{\vec{h}}{a}})$ su šiais teorinės variogramos parametrais: slenkstis $c = 10$, slenksčio nuotolis $a = 1.5$, grynuolis $c_0 = 0$. Remiantis šiuo modeliu, skaičiuojamos ir įrašomos į lentelę žinomų taškų porų pusdispersės $\gamma(\vec{s}_i, \vec{s}_j)$ bei pusdispersės tarp žinomų taškų ir skaičiuojamo taško $\gamma(\vec{s}_i, \vec{s}_0)$, o taip pat atitinkami atstumai.

Taškų poros	Atstumai tarp išmatuotų taškų d_{ij}	Atstumai tarp išmatuotų taškų ir skaičiuojamo taško d_{i0}	Pusdispersės tarp išmatuotų taškų $\gamma(\vec{s}_i, \vec{s}_j)$	Pusdispersės tarp išmatuotų taškų ir skaičiuojamo taško $\gamma(\vec{s}_i, \vec{s}_0)$
(1,2)	4,472		9,493	
(1,3)	6,708		9,886	
(1,4)		4,243		9,409
(2,3)	4,123		9,360	
(2,4)		1,414		6,105
(3,4)		3		8,647

Normaliojo kriginio su *eksponentiniu* pusdispersės modeliu lygtis galima pavaizduoti šia matrica:

$$\begin{bmatrix} 0 & 9.492 & 9.886 & 1 \\ 9.492 & 0 & 9.360 & 1 \\ 9.886 & 9.360 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \Psi \end{bmatrix} = \begin{bmatrix} 9.409 \\ 6.105 \\ 8.647 \\ 1 \end{bmatrix}$$

Šių lygčių sprendinys bus svoriai λ_i :

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \Psi \end{bmatrix} = \begin{bmatrix} 0.1989525 \\ 0.5321374 \\ 0.2689101 \\ 1.6990883 \end{bmatrix}$$

Sisteminės paklaidos nebuvimo apribojimas tenkinamas –
 $0.1989525 + 0.5321374 + 0.2689101 = 1$. Prognozės paklaidą $\hat{\sigma}$ galima apskaičiuoti taip:

$$\hat{\sigma}^2(\vec{s}_0) = \begin{bmatrix} 9.409 & 6.105 & 8.647 & 1 \end{bmatrix} * \begin{bmatrix} 0.1989525 \\ 0.5321374 \\ 0.2689101 \\ 1.6990883 \end{bmatrix} = 9.145$$

Dabar trumpai panagrinėkime kitas kriginio technikas. Gali pasitaikyti atveju, kai regioninis vidurkis žinomas. Kartais verta daryti prielaidą, kad tinka fizikinis modelis, išreiškiamas žinomu trendu. **Paprastojo kriginio** modelyje reikšmė taške $z(\vec{s}_i) = \mu + \eta(\vec{s}_i)$ yra *žinomo pastovaus dydžio* μ ir erdvėje koreliuotos atsitiktinės dedamosios $\eta(\vec{s}_i)$ suma. Šiuo atveju, kadangi μ yra tiksliai žinomas, žinomose vietose tiksliai žinoma ir $\eta(\vec{s}_i)$ reikšmė. Jei žinoma $\eta(\vec{s}_i)$, autokoreliaciją apskaičiuoti lengviau negu tada, kai $\eta(\vec{s}_i)$ irgi reikia įvertinti. Paprastajame kriginėje naudojamos liekanos (skirtumai tarp modelio ir matavimų), darant prielaidą, kad liekanų trendas lygus nuliui.

Paprastojo kriginio sistemoje *Lagranžo daugiklio nereikia*. Paprastajam kriginui *negalioja apribojimas*, kad svorių suma turi būti 1, tai yra $\sum_{i=1}^n \lambda_i \neq 1$. Tačiau skaičiuojant taško reikšmę nuokrypius nuo šių svorių reikia kompensuoti atsižvelgiant į (žinomą) vidurkį. Paprastojo kriginio lygčių sistema yra:

$$\begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_1) & \dots & \gamma(\vec{s}_1, \vec{s}_n) \\ \dots & \dots & \dots \\ \gamma(\vec{s}_n, \vec{s}_1) & \dots & \gamma(\vec{s}_n, \vec{s}_n) \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_0) \\ \dots \\ \gamma(\vec{s}_n, \vec{s}_0) \end{bmatrix}$$

Paprastojo kriginio metodas taip pat gali įvertinti matavimų paklaidas.

Regioninių kintamųjų teorija gali apimti ir *pirmos eilės nestacionarumus* (t. y., kai yra žymus *trendo* paviršius arba žymiai skiriasi *stratų* vidurkiai). *Vidinio stacionarumo* hipotezei pakanka *vietinio* pirmos eilės stacionarumo, taigi normalųjį kriginą galima taikyti vietinėms kaimynystėms, ir šiais atvejais jis tiks. Tačiau net ir tada bus prarasta naudinga informacija apie erdvinę struktūrą. Jei būtų galima atsižvelgti į *globalųjį* trendą, pagerėtų prognozės ir būtų galima geriau suprasti procesus, dėl kurių atsiranda kintamumas erdvėje.

Specialiai sukurti kriginio metodai gali tuo pat metu modeliuoti ir *globalųjį* trendą arba stratifikaciją, ir *vietinę* erdvinės priklausomybės struktūrą. Vienas tokių metodų yra **universalusis kriginas** – procedūra, į kriginio sistemą įtraukianti globalųjį trendą kaip geografinių koordinatų funkciją. Kitas metodas – **regresinis kriginas**, dar vadinamas „kriginu pašalinus trendą“, kai *trendas* (geografinėje arba objektų erdvėje) ir jo *liekanos* modeliuojamos atskirai.

Universalusis kriginas yra mišrus interpoliavimo metodas, kurio sistemoje *globalus trendas* modeliuojamas kaip geografinių koordinatų funkcija. Šiuo atveju kintamojo z reikšmė vietoje $\vec{s}_i$ išreiškiama kaip visos srities *nestacionaraus trendo* $\mu(\vec{s}_i)$ ir *erdvėje koreliuotos atsitiktinės dedamosios* $\eta(\vec{s}_i)$ suma:

$z(\vec{s}_i) = \mu(\vec{s}_i) + \eta(\vec{s}_i)$, kur $\mu(\vec{s}_i)$ nėra pastovus dydis kaip normaliajame kriginėje, bet yra *determinuotoji* padėties (geografinėje erdvėje) funkcija (t. y. *globalus trendas*). Šis trendas modeliuojamas kaip tiesinė P laipsnio žinomų *bazinių funkcijų* $f_j(\vec{s}_i)$ (pvz., globalių daugianarių) ir P nežinomų pastovių koeficientų (modelio parametru) β_j funkcija:

$$z(\vec{s}_i) = \sum_{j=0}^P \beta_j f_j(\vec{s}_i) + \eta(\vec{s}_i).$$
 $\eta(\vec{s}_i)$ yra *erdvėje koreliuota paklaida*, kuri modeliuojama taip, kaip

ir anksčiau – pagal variogramą, tik dabar – *pagal liekanas, gautas pašalinus globalųjį trendą*. Universaliojo kriginio metodu taško reikšmė *prognozuojama* taip pat kaip ir

normaliojo kriginio metodu – $z(\vec{s}_0) = \sum_{i=1}^n \lambda_i z(\vec{s}_i)$, tik kiekvieno išmatuoto taško svoriuose λ_i įvertintas ir *globalus trendas*, ir vietiniai reiškiniai.

Tiesinio slinkio (pirmo laipsnio daugianario) atveju *bazinių funkcijų* pavyzdžiai yra $f_0(\vec{s}_i) = 1$, $f_1(\vec{s}_i) = x$, $f_2(\vec{s}_i) = y$, kur x yra taško abscisė, o y – ordinatė. *Kvadratinio* slinkio atveju būtų ir antro laipsnio nariai – $f_3(\vec{s}_i) = x^2$, $f_4(\vec{s}_i) = xy$ ir $f_5(\vec{s}_i) = y^2$. Atkreipkite dėmesį, kad $f_0(\vec{s}_i) = 1$ įvertina globalų vidurkį – taip pat, kaip ir normaliojo kriginio atveju.

Universaliojo kriginio *sisteminės paklaidos nebuvimo* sąlyga formuluojama atsižvelgiant į

trendą ir bendrą vidurkį (kaip normaliojo kriginio): $\sum_{i=0}^P \lambda_i f_j(\vec{s}_i) = f_j(\vec{s}_0)$ ir $\sum_{i=1}^n \lambda_i = 1$.

Kiekviename taške visų *funkcijų* tikėtina reikšmė turi būti lygi funkcijos reikšmei šiame taške. Pirmoji iš jų, kaip ir normaliajame kriginėje, yra bendras vidurkis.

Universaliojo kriginio pusdispersės $\gamma(\vec{s}_i, \vec{s}_j)$ priklauso ne nuo pradinių duomenų, o nuo *liekanų*, nes atsitiktinė erdvinės struktūros dedamoji susijusi tik su šiomis liekanomis. Pritaikytas variogramos modelis gaunamas trimis etapais:

- skaičiuojamas geriausiai tinkantis trendo paviršius, kuris ir bus naudojamas universaliam kriginui;
- trendas atimamas iš išmatuotų reikšmių ir gaunamos liekanos;
- modeliuojama *liekanų* variograma.

Liekanų modelio *parametrai* paprastai labai skiriasi nuo pradinės variogramos modelio parametrų. Šio modelio slenkstis dažnai žemesnis, o slenkščio atstumas mažesnis, nes atėmus globalųjį trendą, pašalinama dalis kintamumo ir tolimosios struktūros.

Universaliojo kriginio sistemoje sprendžiamos šios lygtys:

$$\begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_1) & \dots & \gamma(\vec{s}_1, \vec{s}_n) & 1 & f_j(\vec{s}_1) & \dots & f_j(\vec{s}_1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma(\vec{s}_n, \vec{s}_1) & \dots & \gamma(\vec{s}_n, \vec{s}_n) & 1 & f_j(\vec{s}_n) & \dots & f_j(\vec{s}_n) \\ 1 & \dots & 1 & 0 & 0 & 0 & 0 \\ f_1(\vec{s}_1) & \dots & f_1(\vec{s}_n) & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ f_p(\vec{s}_1) & \dots & f_p(\vec{s}_n) & 0 & 0 & \dots & 0 \end{bmatrix} * \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \\ \Psi_0 \\ \Psi_1 \\ \dots \\ \Psi_p \end{bmatrix} = \begin{bmatrix} \gamma(\vec{s}_1, \vec{s}_0) \\ \dots \\ \gamma(\vec{s}_n, \vec{s}_0) \\ 1 \\ f_1(\vec{s}_0) \\ \dots \\ f_p(\vec{s}_0) \end{bmatrix}$$

Globalus universalusis kriginas kiekvieną tašką gali skaičiuoti pagal *visų* išmatuotų taškų reikšmes. Tada rezultatai gaunami tokie patys, kaip **regresinio kriginio** metodu. Toks globalus kintamumas tinka tada, kai trendas pasireiškia visoje srityje. Universalusis kriginas gali būti ir *vietinis*, kai taško reikšmė prognozuojama pagal išmatuotus *kaimyninius* taškus. *Vietinis* universaliojo kriginio kintamumas gali atsižvelgti į kintamą srities trendo paviršių, nes kiekviename taške šis paviršius skaičiuojamas iš naujo. Panašiai kaip *paprastojo kriginio* atveju, jei *trendas žinomas*, galima naudoti „**paprastąjį**“ universaliojo kriginio variantą.

Regresinio kriginio atveju $\eta(\vec{s}_i)$ galima gauti atėmus P laipsnio daugianarį iš pradinių duomenų, o $\eta(\vec{s}_i)$ laikomos atsitiktinėmis. Visų $\eta(\vec{s}_i)$ vidurkis lygus 0. Dabar autokovariacija iš esmės modeliuojama pagal atsitiktines paklaidas $\eta(\vec{s}_i)$. Regresinį kriginą galima laikyti vienu iš polinominės regresijos tipų. Tačiau šiuo atveju nelaikoma, kad paklaidos $\eta(\vec{s}_i)$ yra nepriklausomos – jos modeliuojamos lyg būtų autokoreliuotos. Universaliojo ir regresinio kriginio metodai taip pat gali įvertinti matavimų paklaidas.

Regresinis kriginas atliekamas šiais etapais:

- skaičiuojamas *trendas* ir gaunama jo tiesinio modelio prognozės paklaida;
- atėmus trendą, gaunamos žinomų taškų *liekanos*;
- modeliuojamos kriginio liekanos – tai galima atlikti paprastojo kriginio atveju, kai žinomas liekanų vidurkis yra 0 ir įmanoma nustatyti liekanų prognozės paklaidą;
- vėl pridėjus trendą prie sumodeliuotų liekanų, gaunami nežinomų taškų įverčiai;
- visa paklaida kiekviename taške gaunama sudėjus dvi apskaičiuotas prognozės dispersijas (tiesinio modelio prognozės paklaidą ir liekanų prognozės paklaidą).

Gali būti, kad matuojami duomenys yra *dvejetainiai* (reikšmės yra 0 ir 1) arba tolydus kintamasis perklasifikuotas į dvejetainį pagal tam tikrą schemą – pvz., jei reikšmės

didesnės už ribą, jos tampa 1, jei mažesnės – 0. Pavyzdžiui, paviršių galima klasifikuoti kaip sausumą (1) ir vandenį (0). **Indikatorinio kriginio** modelis yra $z(\vec{s}_i) = \mu(\vec{s}_i) + I(\vec{s}_i)$, kur $\mu(\vec{s}_i)$ yra nežinomas pastovus dydis, o $I(\vec{s}_i)$ – *dvejietinis kintamasis*. Indikatorinio kriginio skaičiavimai visiškai tokie patys kaip normaliojo, tik atliekami su dvejetainiais kintamaisiais. Interpoliuotos reikšmės bus tarp 0 ir vieno 1 – indikatorinio kriginio prognozę galima interpretuoti kaip tikimybę, kad kintamojo reikšmė yra 1 arba kad jis priklauso klasei, kuri nurodyta reikšme 1.

Regioninių kintamųjų kriginio teorijos išplėtimas iki kelių kintamųjų, turinčių *daugiamatę kryžminę erdvinę koreliaciją* bei atskiras *vienmatis erdvinės autokoreliacijas*, vadinamas **koregionalizacija**. **Kokrigingas** yra metodas, *ieškomo kintamojo prognozei* patikslinti naudojantis pagalbinę koregionalizuotų (*lydinčiųjų*) kintamųjų informaciją.

Koregionalizacijos esmė ta, kad vieną kintamąjį lemiantis procesas yra toks pats kaip kitus kintamuosius aprašantys procesai arba bent jau susijęs su jais. Pavyzdžiui, sunkiųjų metalų pasiskirstymas dirvožemyje gali būti susijęs su užterštumo ir vandens pH pasiskirstymu bei aukščiu virš jūros lygio. Visi nagrinėjami kintamieji turi būti *regioniniai*; be to, jei jie visi susiję geografinėje erdvėje, jie *koregionalizuojami*.

Pavyzdžiui, išmatuotos tik kelios *ieškomo kintamojo* reikšmės, o papildomus matavimus atlikti yra labai brangu. Bet jei žinome daugiau *antro kintamojo* reikšmių ir antras kintamasis *koreliuoja* su ieškomu kintamuoju, jis tampa *lydinčiuoju kintamuoju*. Šio antro kintamojo reikšmių išmatuota daug, nes, pavyzdžiui, matavimai paprasti ir pigūs. Dažniausiai *išmatuotų lydinčiojo kintamojo reikšmių būna daugiau* (t. y. ieškomas kintamasis išmatuotas ne visuose taškuose).

Taigi, koregionalizacijos teorijoje naudojamos dviejų tipų variogramos: pirmoji yra *tiesioginė* – atskira vieno regioninio kintamojo variograma, antroji yra *kryžminė* – regioninių kintamųjų poros variograma. *Kryžminė empirinė* variograma skaičiuojama atstumams,

kurie yra $\vec{h}$ kartotiniai, pagal formulę $\bar{\gamma}_{1,2}(\vec{h}) = \frac{1}{2m(\vec{h})} \sum_{k=1}^{m(\vec{h})} [Z_1(\vec{s}_i) - Z_1(\vec{s}_j)] * [Z_2(\vec{s}_i) - Z_2(\vec{s}_j)]$, kur

$Z_1(\vec{s}_i)$ yra *ieškomas* arba pagrindinis dominantis kintamasis, o $Z_2(\vec{s}_i)$ – *lydintysis* arba koregionalizuotas kintamasis. **Kryžminės variogramos** gali vaizduoti arba teigiamą, arba neigiamą erdvinę koreliaciją.

Tiesioginės ir kryžminės variogramos turi būti modeliuojamos kartu, laikantis tam tikrų apribojimų, užtikrinančių, kad kokrigingo lygčių sistemą bus įmanoma išspręsti. Todėl kokrigingo metode naudojama informacija apie *kelis kintamųjų tipus*. Bendrieji **normaliojo kokrigingo** modeliai yra:

$$z_1(\vec{s}_i) = \mu_1(\vec{s}_i) + \eta_1(\vec{s}_i)$$

$$z_2(\vec{s}_i) = \mu_2(\vec{s}_i) + \eta_2(\vec{s}_i)$$

kur $\mu_1(\vec{s}_i)$ ir $\mu_2(\vec{s}_i)$ yra *nežinomi pastovūs dydžiai*. Atsitiktinės paklaidos yra dviejų tipų – $\eta_1(\vec{s}_i)$ ir $\eta_2(\vec{s}_i)$, taigi yra kiekvienos jų autokoreliacija ir kryžminė koreliacija tarp jų. Kaip ir normaliuoju kriginu, normaliuoju kokrigingu bandoma prognozuoti $z_1(\vec{s}_i)$, bet prognozę bandoma pagerinti remiantis lydinčiojo kintamojo $\{z_2(\vec{s}_i)\}$ informacija.

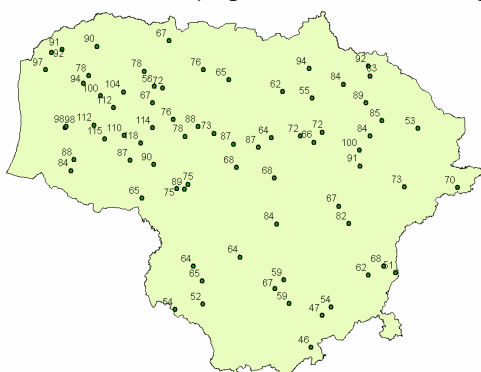
Ieškomas kintamasis yra $z_1(\vec{s}_i)$, o prognozei pagerinti naudojama ir $z_1(\vec{s}_i)$ autokoreliacija, ir kryžminė koreliacija tarp $z_1(\vec{s}_i)$ ir kito tipo kintamojo $z_2(\vec{s}_i)$. Lydinčiojo kintamojo $z_2(\vec{s}_i)$ informacija naudojama kaip pagalbinė, bet ją įvertinti daug sudėtingiau, nes reikia skaičiuoti kiekvieno kintamojo autokoreliaciją ir kryžmines koreliacijas. Teoriškai kokrigingo rezultatai negali būti blogesni už krigingo, nes jei nėra kryžminės koreliacijos, jis pagrįstas tik $z_1(\vec{s}_i)$ autokoreliacija. Tačiau praktiškai dėl nežinomų autokoreliacijos parametrų įvertinimo gali padidėti kintamumas ir sumažėti prognozės tikslumas.

Normaliojo kokrigingo interpoliavimo formulė yra svertinė duomenų suma – $z(\vec{s}_0) = \sum_{i=1}^n \lambda_i Z_1(\vec{s}_i) + \sum_{i=1}^n \omega_i Z_2(\vec{s}_i)$, kur $Z_1(\vec{s}_i)$ yra išmatuota ieškomo kintamojo reikšmė, λ_i – nežinomas išmatuotos ieškomo kintamojo reikšmės svoris, $Z_2(\vec{s}_i)$ – išmatuota lydinčiojo kintamojo reikšmė, ω_i – nežinomas išmatuotos lydinčiojo kintamojo reikšmės svoris, $z(\vec{s}_0)$ – prognozė vietoje $\vec{s}_0$, n – grupės matavimų skaičius.

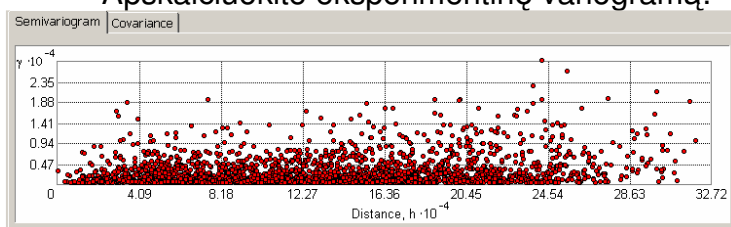
Kiti kokrigingo metodai, įskaitant universalųjį kokrigingą, paprastąjį kokrigingą, indikatorinį kokrigingą, yra anksčiau aprašytų metodų apibendrinimai keliems duomenų rinkiniams. Kokrigingo metodai gali įvertinti matavimų paklaidas taip pat, kaip ir atitinkami krigingo metodai (normalusis krigingas, paprastasis krigingas ir universalusis krigingas).

Krigingo metodus galima sudalinti į šiuos optimalaus erdvinės prognozės modeliavimo etapus:

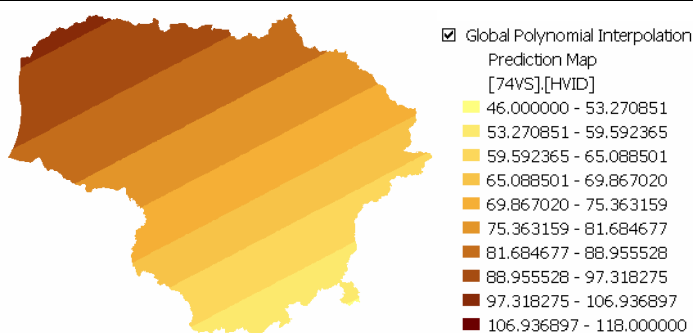
- Išmatuokite reiškinį (pvz., pavasario potvynio vidutinis daugiametis nuotėkio aukštis), geriausia – skirtinga skiriamąja geba:



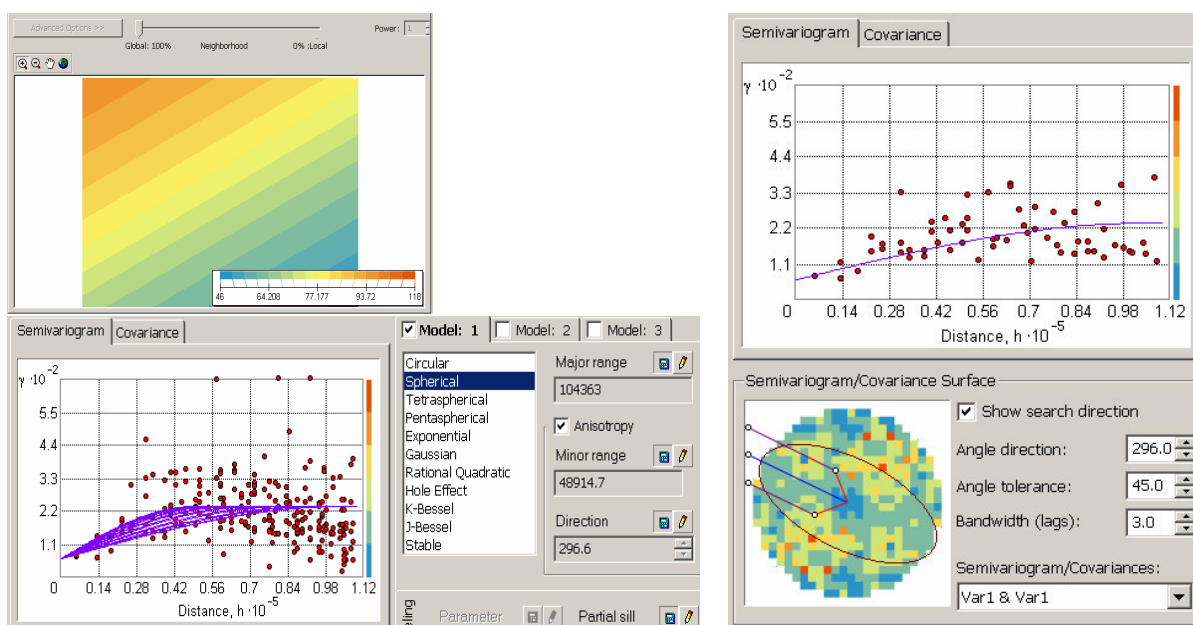
- Apskaičiuokite eksperimentinę variogramą:



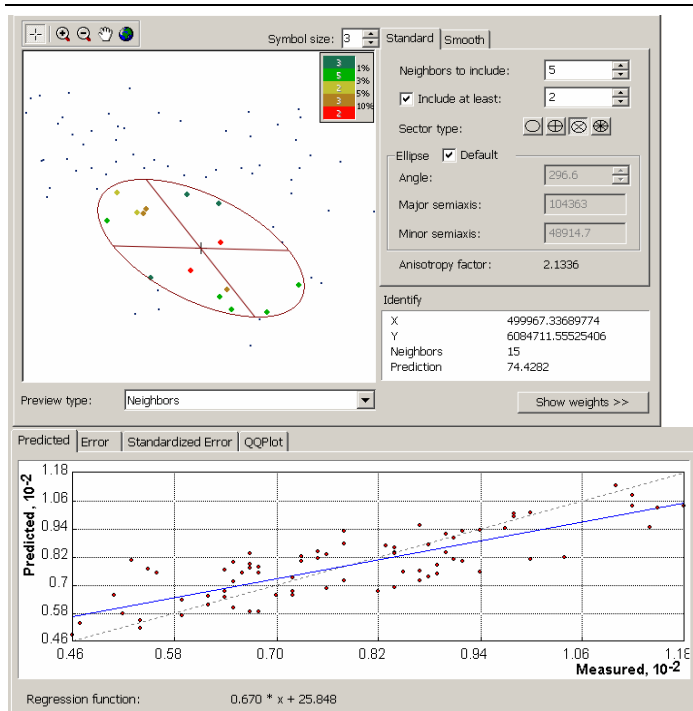
- Patikrinkite, ar yra trendas (pvz., pavaizduokite interpoliavę naudodami globaliojo polinomo metodą). Jei trendas yra, nustatykite trendo paviršiaus modelį (šio atveju yra tiesinis trendas šiaurės vakarų kryptimi):



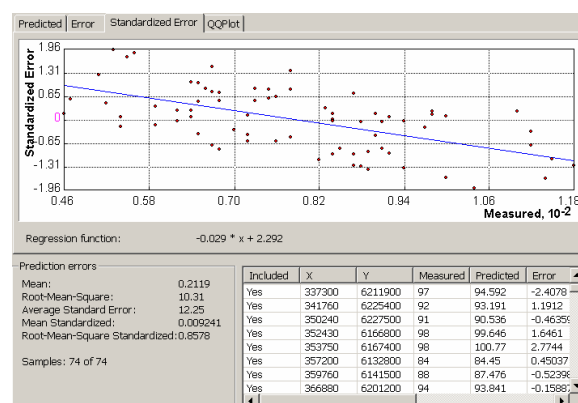
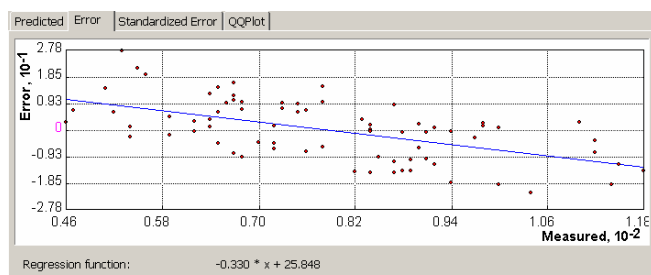
- Sumodeliuokite liekanų variogramą $\eta(\bar{s}_i) = z(\bar{s}_i) - \mu(\bar{s}_i)$ naudodami vieną arba kelias leistinas funkcijas. $\eta(\bar{s}_i)$ gaunama iš pradinių duomenų atėmus pirmo laipsnio daugianarį (kairysis viršutinis paveikslėlis). Kairiojoje variogramoje kiekviena mėlyna kreivė vaizduoja pritaikytą konkrečios paieškos krypties (šiaurės, šiaurės vakarų ir t. t.) modelį (anizotropijos modeliai). Dešiniojoje variogramoje pavaizduotas pritaikytas šiaurės vakarų krypties modelis (mėlyna kreivė), atitinkantis pagrindinės aprėpties kryptį.



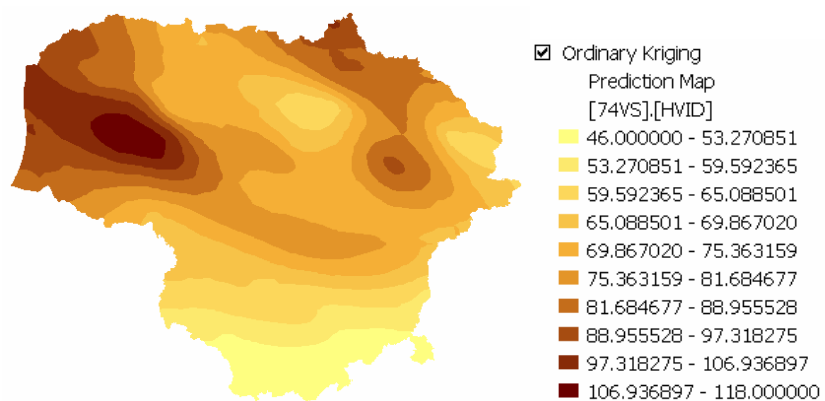
- Kai kurios krigingo sistemos, kuriose erdvinės priklausomybės modeliai yra variogramos, gali prognozuoti reikšmes kiekviename taške pagal variogramos ir trendo modelį. Dažnai skaičiuojamos visų taisyklingo tinklėlio (pvz., rastrinio žemėlapis) taškų reikšmės. Jei krigingas vietinis, sprendžiama atskira lygčių sistema kiekvienam taškui $\bar{s}_0$, naudojant pusdisperses aplink tą tašką (*vietinėje kaimynystėje*); kiekvienam prognozuojamam taškui gaunami skirtingi svorių rinkiniai λ_i (kairysis paveikslėlis). Apskaičiuotų pritaikytos funkcijos reikšmių grafikas vaizduojamas per išsklaidytus išmatuotus taškus einančia mėlyna linija, kurios formulė užrašoma po grafiku (dešinysis paveikslėlis). Jei duomenų autokoreliacijos nėra, visos prognozės bus vienodos, kitaip tariant, kiekviena prognozė bus lygi žinomų reikšmių vidurkiui – mėlyna linija bus horizontali. Jei autokoreliacija yra ir krigingo modelis geras, mėlyna linija bus arčiau juodos brūkšninės linijos.



- Apskaičiuojamos kiekvienos prognozės paklaidos; jos priklauso tik nuo išmatuotų taškų vietų ir *nepriklauso* nuo duomenų reikšmių. Paklaidų grafikas yra toks pat kaip ir prognozių grafikas, tik iš apskaičiuotų reikšmių atimtos išmatuotos reikšmės (kairysis paveikslėlis). Standartinių paklaidų grafike vaizduojami apskaičiuotų ir išmatuotų verčių skirtumai, padalinti iš įvertintų kriginio prognozės paklaidų $\hat{\sigma}$ bei kitų išvestinių standartinių paklaidų (dešinysis paveikslėlis).



- *Apskaičiuotos reikšmės* pavaizduojamos žemėlapyje:



4.5 Modelio patikra

Kad ir koks būtų interpoliavimo metodas, norisi žinoti, ar geri bus jo rezultatai. Todėl sukurtą modelį reikia patvirtinti faktais. Su tuo susijusi **modelio patikros** samprata ir metodika. Modelio patikra, įskaitant modelio *kalibravimą*, atliekama kuriant modelį.

Ar modelis atitinka tikrovę, galima patikrinti šiais būdais:

- Nepriklausomais duomenų ir modelių tinkamumo **matavimais** – kai kuriuos iš jų galima pavaizduoti paviršiumi.
- Atskiru **patikros** duomenų rinkiniu – turimus duomenis padalinti į dvi dalis – mokomąją ir bandomąją, ir palyginti pagal mokomąjį rinkinį apskaičiuotas reikšmes nurodytose vietose su bandomojo rinkinio reikšmėmis.
- **Kryžminės patikros** būdu, naudojant kalibravimo duomenų rinkinius.
- Sukurti daug interpoliavimo paviršių ir **lyginti** gautus rezultatus.

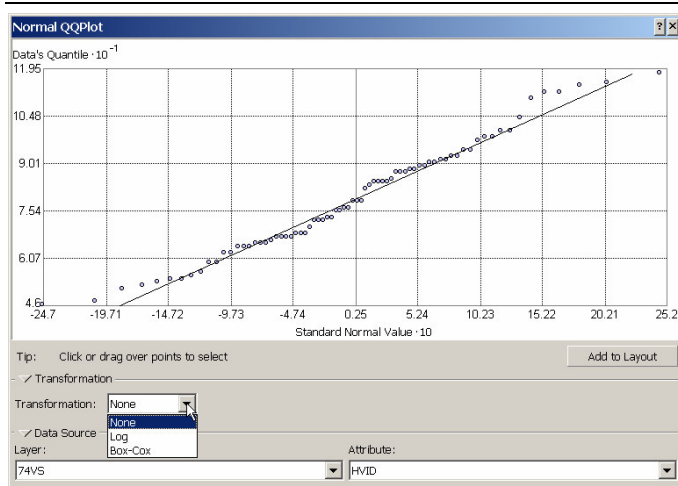
4.5.1 Duomenų tinkamumas

Krigingas pagrįstas duomenų modeliavimu. Kad būtų galima teisingai interpretuoti rezultatus, erdviniam modeliavimui naudojamus duomenis reikia iš anksto patikrinti.

Duomenų tyrimo etapu, paprastai atliekamu prieš interpoliavimą, reikia patikrinti duomenų normalumą, išskirtis, globalųjį trendą, stacionarumą, autokoreliaciją erdvėje ir t. t. Kai kuriems krigingo ir kokrigingo metodams (pvz., normaliajam krigingui) reikia, kad duomenų skirstinys būtų normalusis, todėl reikia patikrinti duomenų normalumą. Jei duomenys pasiskirstę ne pagal normalųjį skirstinį, duomenis gali tekti transformuoti normaliųjų taškų transformacijomis.

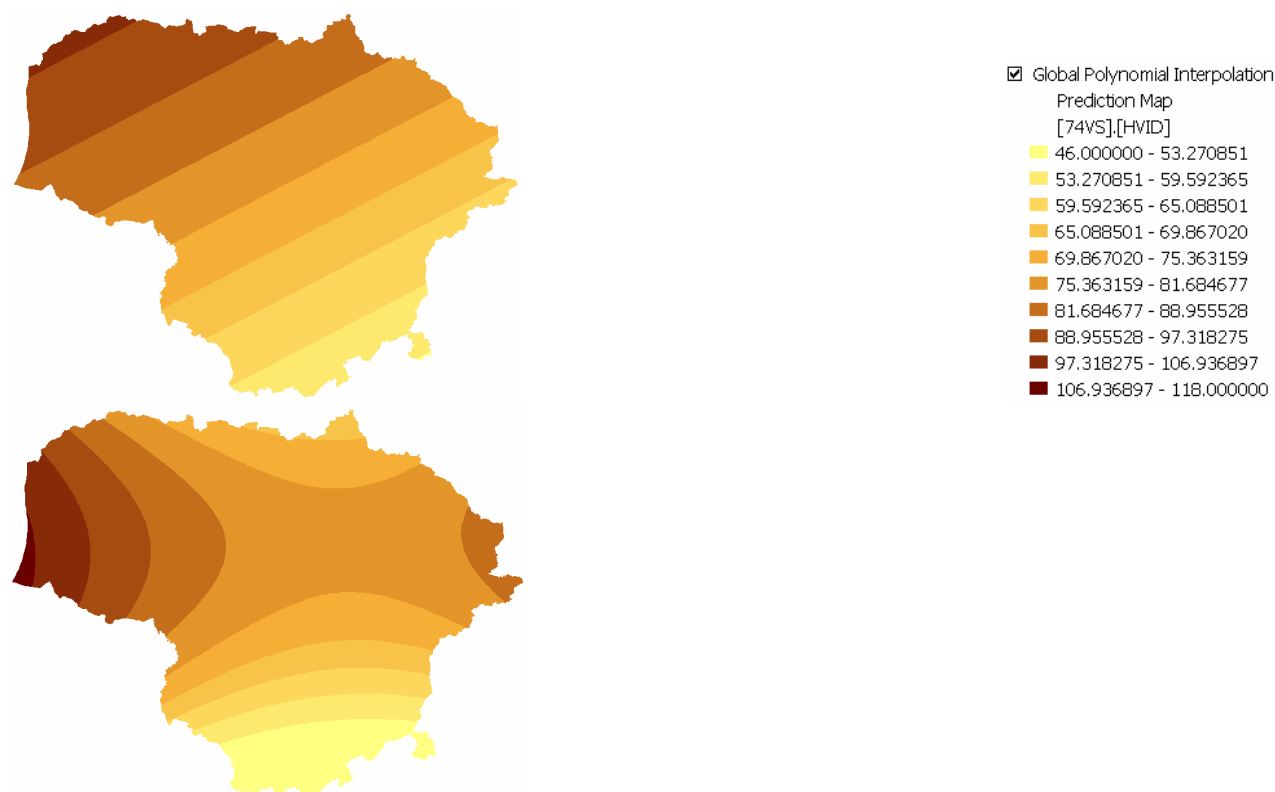
Ar duomenys *normaliai pasiskirstę*, galima nustatyti iš histogramos. Normaliojo skirstinio histograma yra vienviršūnio varpo formos, vidutinė ir vidurinė (*median*) reikšmės lyra labai artimos, o ekscesas yra beveik 3. Normalumą galima patikrinti ir skaitiniais kriterijais (pvz., Šapiro ir Vilkso arba Andersono).

Be to, galima naudoti *normaliojo skirstinio kvantilių grafiką* (*normal QQPlot*), kurio x ašyje atidedami standartinio normaliojo skirstinio kvantiliai, o y ašyje – įvesties duomenų rinkinio kvantiliai (4.26 pav.). Kvantilių grafikais galima grafiškai palyginti du skirstinius. Jei skirstiniai tapatūs, normaliojo skirstinio kvantilių grafikas bus tiesė. Taigi, vienmatį normalumą galime įvertinti normaliojo skirstinio kvantilių grafike lygindami šią tiesę su išmatuotų taškų skirstiniu. Jei išmatuotų taškų skirstinys asimetris (t. y. labai skiriasi nuo normaliojo), taškai bus nutolę nuo šios tiesės.



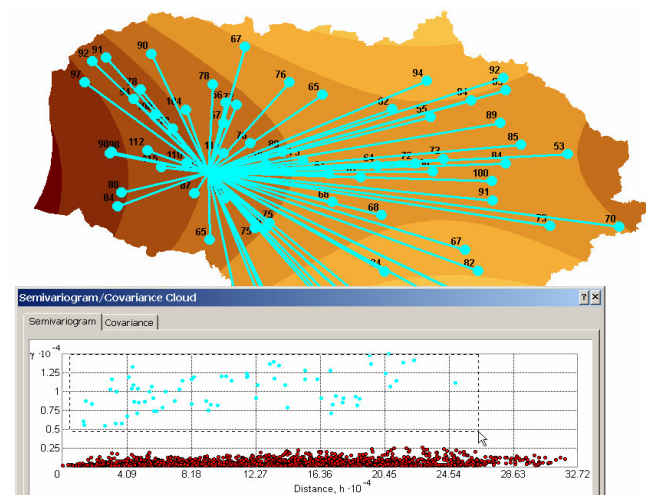
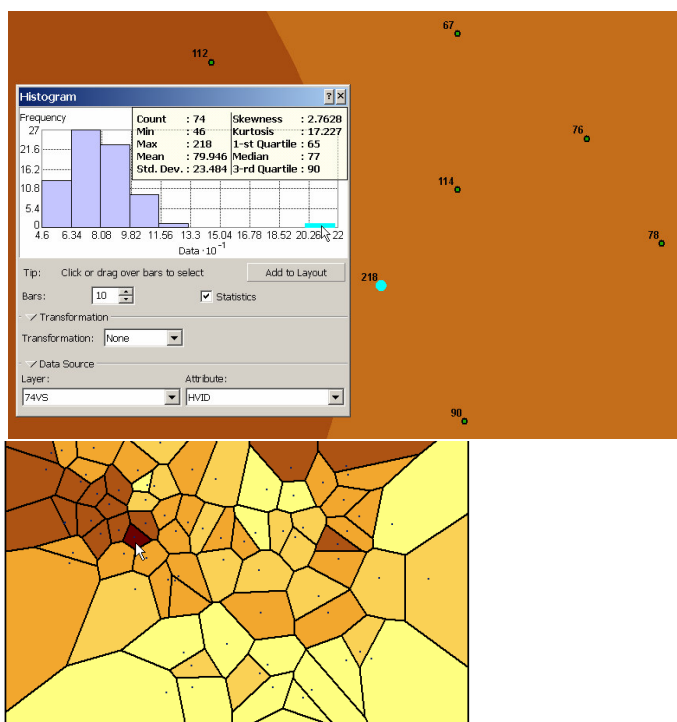
4.26 pav. Ilgalaikių vidutinių vandens lygio reikšmių normaliojo skirstinio kvantilių grafikas ir ESRI Geostatistical Analyst transformacijos parinktys. Daugiau informacijos apie transformaciją rasite <http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Box-Cox%2C Arcsine%2C and Log transformations> ir <http://www.spatialanalysisonline.com/output/html/Datatransformsandbacktransforms.html>.

Vizualiai patikrinti, ar duomenyse pasireiškia tiesinis arba kitoks trendas, galima globalaus polinominio interpoliavimo metodais. Informacijos apie globalias duomenų tendencijas gali suteikti ir duomenų reikšmių skersinių pjūvių grafikai.



4.27 pav. Pavasario potvynio vidutinio daugiamečio nuotėkio aukščio reikšmės turi tiesinę (pirmo laipsnio) tendenciją (žemėlapis kairėje). Aiškios antro laipsnio tendencijos nematyti (žemėlapis dešinėje)

Duomenų išskirtis (*data outliers*) – staigius reikšmių pokyčius, galinčius atsirasti dėl tikrų reiškinių anomalijų arba matavimo klaidų, galima nustatyti iš histogramos (taškai skirstinio gale), pusvariogramės (nuo atstumo nepriklausančių didelių reikšmių taškų poros pusvariogramės taškų debesyje) ir Voronojaus poligonų (didelis skirtumas tarp kaimynų) – 4.28 pav. Apie Voronojaus poligonus galite paskaityti http://en.wikipedia.org/wiki/Voronoi_diagram.



4.28 pav. Išskyros pavyzdys – taškas, kurio reikšmė 218

Paviršių anizotropiją galima tirti kryptinėmis pusvariogramėmis, aptartomis 1.4 skyriuje.

4.5.2 Nepriklausomi modelio tinkamumo matai

Interpoliavimo rezultatus galima vertinti įvairiais nepriklausomais tinkamumo matais. Štai keli iš jų:

Absoliučių reikšmių arba liekanų *vidutinis kvadratinė paklaida* (**VKP**, angl. RMSE - *root-mean-squared error*) – patikros duomenų rinkinio tikrosios reikšmės lyginamos su modelio

apskaičiuotomis reikšmėmis: $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z(\vec{s}_i) - Z(\vec{s}_i))^2}$, kur $z(\vec{s}_i)$ yra prognozuojama reikšmė, o $Z(\vec{s}_i)$ – išmatuota reikšmė.

Sisteminė paklaida arba *patikros duomenų rinkinio vidutinis apskaičiuotos reikšmės nuokrypis* (**VN**, angl. ME - *mean error*) nuo tikrojo vidurkio: $ME = \frac{1}{n} \sum_{i=1}^n (z(\vec{s}_i) - Z(\vec{s}_i))$

Krigingo prognozės paklaida $\hat{\sigma}$ konkrečiame taške arba krigingo dispersija konkrečiame taške: $\hat{\sigma}^2(\vec{s}_0) = \mathbf{b}^T \boldsymbol{\lambda}$.

Vidutinė krigingo prognozės paklaida: $\sqrt{\frac{1}{n} \sum_{i=1}^n \hat{\sigma}(\vec{s}_i)}$.

Vidutinė standartinė prognozės paklaida: $\frac{1}{n} \sum_{i=1}^n (z(\vec{s}_i) - Z(\vec{s}_i)) / \hat{\sigma}(\vec{s}_i)$.

Vidutinė kvadratinė standartinė prognozės paklaida: $\sqrt{\frac{1}{n} \sum_{i=1}^n ((z(\vec{s}_i) - Z(\vec{s}_i)) / \hat{\sigma}(\vec{s}_i))^2}$.

Prediction errors	
Mean:	0.2119
Root-Mean-Square:	10.31
Average Standard Error:	12.25
Mean Standardized:	0.009241
Root-Mean-Square Standardized:	0.8578
Samples: 74 of 74	

4.29 pav. Gauto pavasario potvynio vidutinio daugiamečio nuotėkio aukščio paviršiaus reikšmių prognozės paklaidos

Lyginant modelius, reikia ieškoti tokio, kuris tenkintų šias sąlygas:

- Vidutinė kvadratinė paklaida mažesnė
- Sistemine paklaida artima nuliui
- Vidutinė standartinė prognozės paklaida arčiausiai nulio
- Vidutinė prognozės paklaida artimiausia vidutinei kvadratinei paklaidai
- Vidutinė kvadratinė standartinė prognozės paklaida artimiausia vienetui

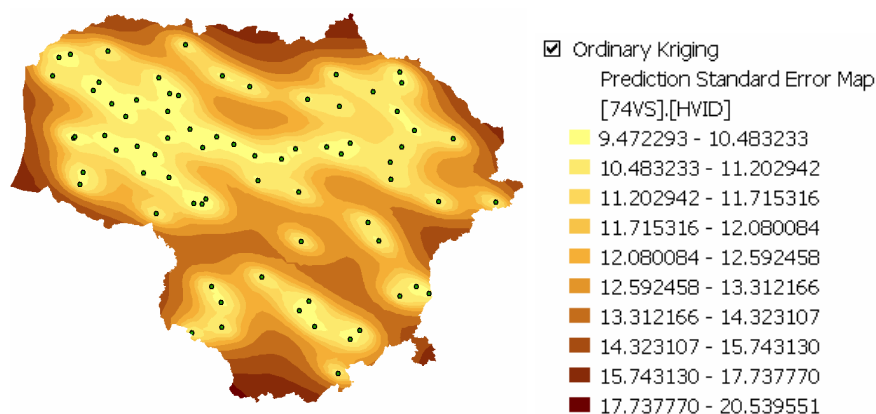
Krigingo prognozės paklaidą $\hat{\sigma}$ galima įvertinti kiekviename skaičiuojamame taške. Todėl naudojant stochastinį metodą, tokį kaip krigingas, prognozės paklaidas galima išreikšti skaičiais ir pavaizduoti paklaidų paviršiumi arba žemėlapiu. Determinuotieji interpoliavimo metodai (atvirkščiai proporcingo atstumo, splainų ir pan.) paklaidų neįvertina! Todėl lyginti determinuotuosius interpoliavimo metodus su kitais metodais galima tik pagal VKP.

Bet kaip apskaičiuoti tikslaus interpoliavimo (*exact interpolation*) VKP? Tikslaus interpoliavimo VKP galima apskaičiuoti [kryžminės patikros](#) būdu. Atliekant kryžminę patikrą, pašalinamas vienas išmatuotas taškas, tada interpoliuotas paviršius skaičiuojamas be šio taško ir palyginamas su tikrąja pašalinto taško reikšme. Ši procedūra kartojama visiems išmatuotiems taškams. Kryžminės patikros VKP yra bendroji statistika, įvertinanti prognozuojamo paviršiaus paklaidą. Išsamiau kryžminė patikra aprašyta kitame šios dalies skyriuje.

Kol kas kūrėme **prognozių** žemėlapius, dar vadinamus interpoliuotais žemėlapiais. Prognozių žemėlapiuose vaizduojamos interpoliuotos reikšmės. **Paklaidų** žemėlapiuose (pvz., krigingo prognozės paklaidų žemėlapiuose) vaizduojamos reikšmės lygios

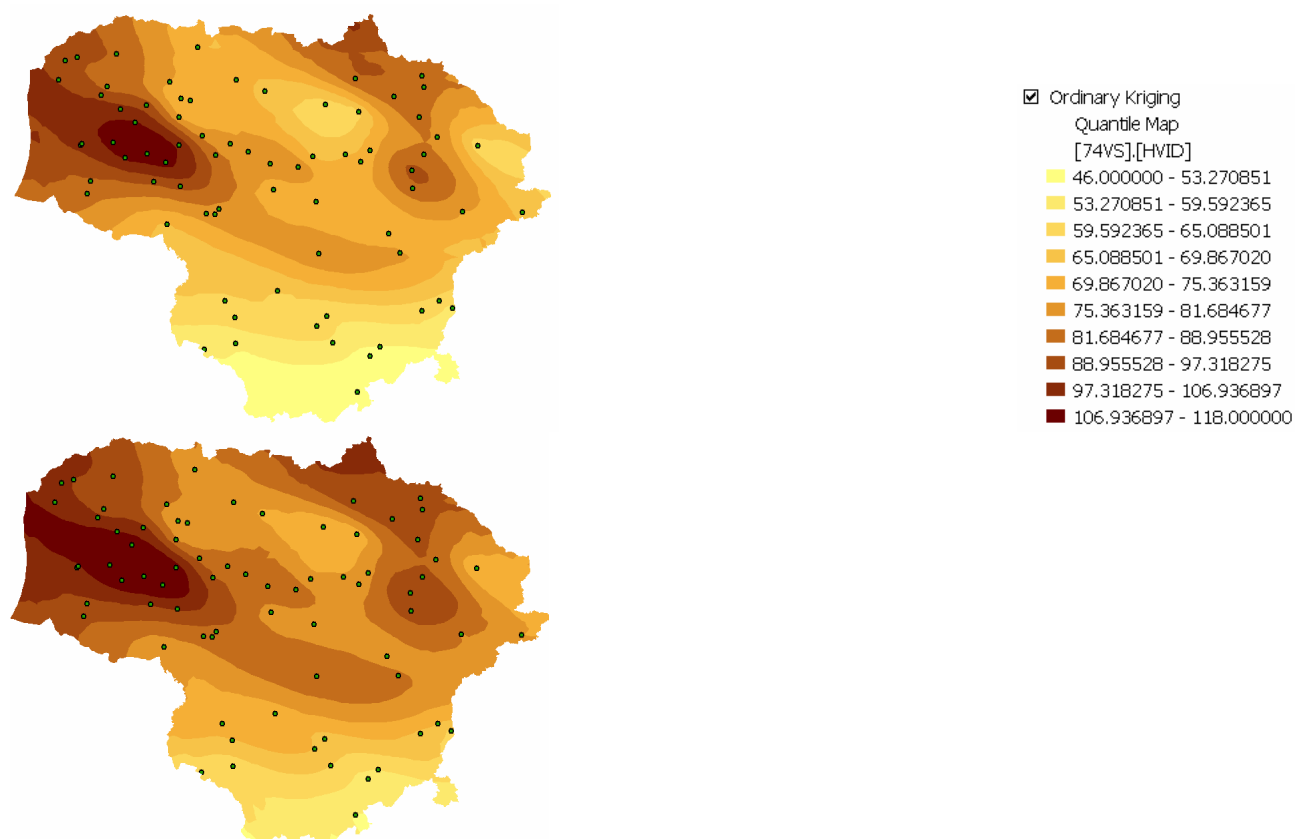
kvadratinei šakniai iš krigingo prognozės dispersijos ($\hat{\sigma}^2$ įverčio) – 4.30 pav. Paklaidų žemėlapiai vaizduoja prognozės neapibrėžtumą. Jei duomenų skirstinys yra normalusis, 95 proc. atvejų tikroji reikšmė bus ± 2 prognozės paklaidų intervale nuo vidurkio.

Galima ne tik prognozuoti reikšmes, bet ir įvertinti prognozių kintamumą tikrųjų reikšmių pagrindu. Jei *vidutinė krigingo prognozės paklaida* didesnė už *vidutinę kvadratinę paklaidą*, vadinasi, prognozės kintamumas yra padidintas; jei *krigingo prognozės paklaida* mažesnė už *vidutinę kvadratinę paklaidą* – prognozės kintamumas yra per sumažinamas.



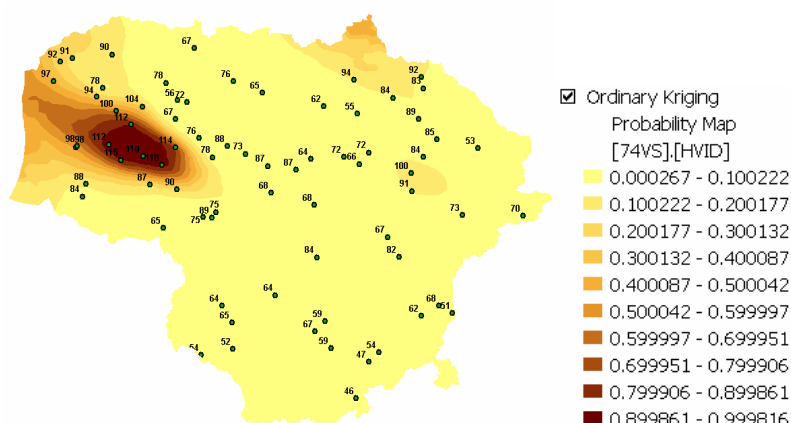
4.30 pav. Pavasario potvynio vidutinio daugiamečio nuotėkio aukščio reikšmių įvertinimo paklaidų žemėlapis. Paklaidos didesnės ten, kur matavimai retesni.

Iš paklaidų žemėlapių galima gauti du kitus paklaidų įvertinimo žemėlapius. Tai **kvantilių** (4.31 pav.) ir **tikimybių** (4.32 pav.) žemėlapiai. **Kvantilių** žemėlapių reikšmės atspindi tikrųjų reikšmių viršutinės arba apatinės ribas. Kvantilių žemėlapiai – tai paviršiai reikšmių, kurias prognozuojamos reikšmės viršija (arba kurių neviršija) su *nurodyta* tikimybe. Pavyzdžiui, jei kvantilio tikimybė yra 0,5, žemėlapyje bus vaizduojamos visų žinomų taškų vidurinės prognozuojamos reikšmės (medianos). Jei kvantilio tikimybė 0,75, bus sukurtas toks paviršius, kad tikimybė, jog nežinoma reikšmė yra mažesnė už paviršiaus reikšmę, bus 75 proc., o tikimybė, kad didesnė už paviršiaus reikšmę – 25 proc.



4.31 pav. Apskaičiuotų pavasario potvynio vidutinio daugiamečio nuotėkio aukščio reikšmių kvantilių žemėlapis, kvantilių tikimybės 0,5 ir 0,75

Tikimybų žemėlapis vaizduoja tikimybę, kad tikroji reikšmė bus didesnė (mažesnė) už nustatytą slenkstį. Pavyzdžiui, jei slenkščio reikšmė yra „*daugiau negu 100*“, žemėlapis vaizduos tikimybes, kad reikšmės viršys 100 matavimo vienetų.

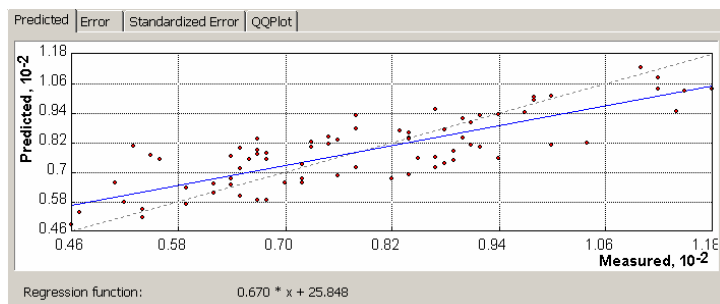


4.32 pav. Apskaičiuotų pavasario potvynio vidutinio daugiamečio nuotėkio aukščio reikšmių tikimybų žemėlapis, slenkstis – daugiau negu 100.

Kad tikimybų ir kvantilių žemėlapis būtų patikimi, duomenų skirstinys turi būti visiškai daugiamačis normalusis skirstinys.

Krigingo prognozės rezultatams tikrinti tinka tik keli grafikų tipai. Vienas iš jų yra sklaidos grafikas (*scatter plot*), kuriame prognozuojamos reikšmės (paveikslėlyje – mėlyna

pritaikyta linija) lyginamos su išmatuotomis reikšmėmis (taškais). Galima tikėtis, kad pritaikyta linija turi sutapti su grafiko įstrižaine (juoda brūkšninė linija). Tačiau linijos nuolydis paprastai mažesnis už vienetą. Viena iš kriginio ypatybių yra tendencija sumažinti dideles prognozuojamas reikšmes ir padidinti mažas, tai yra, glodinti paviršių (4.33 pav.).



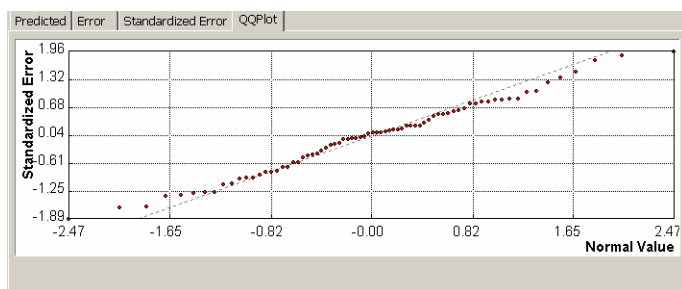
4.33 pav. Prognozuojamų ir išmatuotų reikšmių sklaidos grafikas. Pritaikyta linija pavaizduota mėlyna spalva, linijos formulė pateikta po grafiku

Jei išmatuoti duomenys nekoreliuoja tarpusavyje (nepriklausomi erdvėje), visos prognozės bus vienodos ir lygios išmatuotų reikšmių vidurkiui. Tada mėlyna linija bus horizontali. Jei autokoreliacija yra ir kriginio modelis teisingas, mėlyna linija turėtų būti arčiau juodos brūkšninės linijos. Kuo taškai arčiau įstrižainės, tuo geriau.

Be to, *paklaidų* kvantilių grafiku prognozės paklaidų skirstinį galima palyginti su standartiniu normaliuoju skirstiniu. Šis paklaidų grafikas toks pat kaip prognozės grafikas, tik iš apskaičiuotų reikšmių atimtos išmatuotos reikšmės.

Sudarant *standartinių paklaidų* kvantilių grafiką, iš apskaičiuotų reikšmių atimtos išmatuotos reikšmės, ir skirtumas dalijamas iš apskaičiuotų kriginio paklaidų.

Jei standartinės prognozės paklaidos pasiskirsčiusios normaliai, taškai turi būti arti normalųjį skirstinį vaizduojančios linijos (4.34 pav.). Jei šios paklaidos pasiskirsčiusios normaliai, tai galima naudoti normalumu pagrįstus metodus (pvz., normalųjį kriginą).



4.34 pav. Prognozės standartinių paklaidų kvantilių grafikas. Pavasario potvynio vidutinio daugiamečio nuotėkio aukščio reikšmių skirstinys artimas normaliajam

4.5.3 Modelio patikra poaibiais

Griežčiausias gauto paviršiaus kokybės tikrinimo būdas – palyginti konkrečiose vietose prognozuojamas reikšmes su nepriklausomai išmatuotais kontroliniais taškais. Tikrinimui naudojami du duomenų rinkiniai. Modelis sukuriamas pagal vieną duomenų poaibį, vadinamą *mokomuoju* duomenų rinkiniu, o tikrinamas pagal kitą poaibį – *bandomąjį*

kontrolinių taškų duomenų rinkinį. Vertinimo kriterijus gali būti, pavyzdžiui, iš *mokomųjų* reikšmių gauto modelio prognozių *vidutinė kvadratinė paklaida bandomųjų* reikšmių kontroliniuose taškuose atžvilgiu.

Šiuo atveju atliekant patikrą modelis sukuriamas tik vienam duomenų poaibiui – *mokomajam* duomenų rinkiniui; todėl galutinis modelis, kuris turėtų būti sudarytas atsižvelgiant į visus turimus duomenis, nėra tikrinamas tiesiogiai. Patikros metu nustatoma, ar tinkamas *mokomojo* duomenų rinkinio modelis, pavyzdžiui, ar gerai pasirinktas pusvariogramės modelis, vėlavimo žingsnis ir paieškos kaimynystė. *Mokomojo* duomenų rinkinio modelis naudojamas visam duomenų rinkiniui.

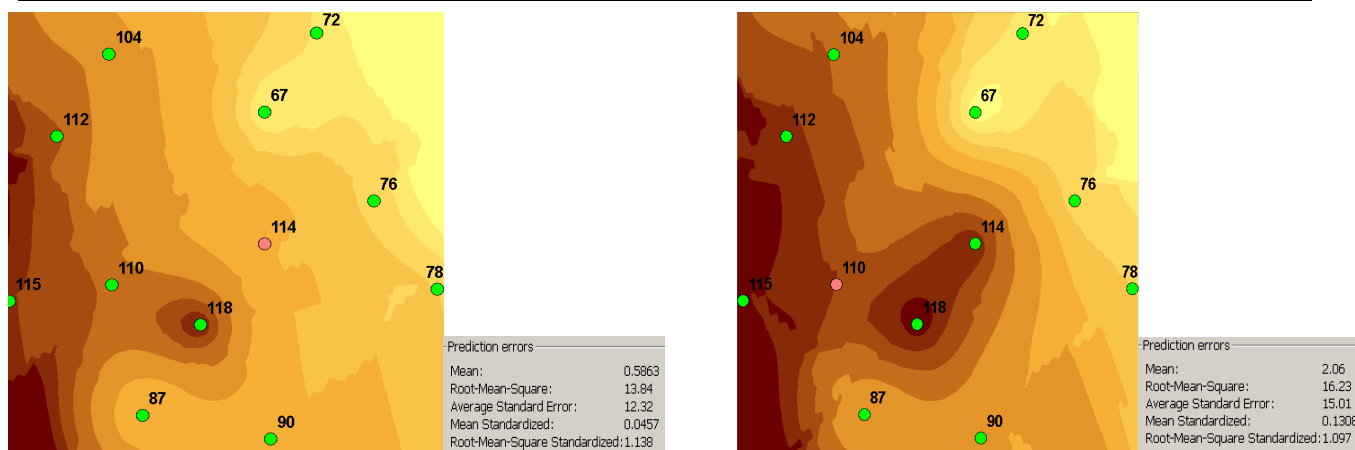
4.5.4 Kryžminė modelio patikra

Jei neturime nepriklausomo duomenų rinkinio, pagal kurį galėtume įvertinti modelį, jį galima patikrinti naudojant tuos pačius *išmatuotus taškus*, iš kurių modelis buvo sudarytas. Taigi, atliekant kryžminę patikrą (*cross-validation*), trendams ir autokoreliacijos modeliams tikrinti naudojami visi išmatuoti duomenys. Po vieną pašalinamas kiekvienas taškas, ir prognozuojama reikšmė tame taške bei apskaičiuojama prognozės paklaida (4.35 pav.). Analizuojant gautas paklaidas (4.36 pav.) galima nustatyti išskyras, o po kryžminės patikros kai kurias vietas atmesti kaip neįprastas ir patikslinti tendencijos bei autokoreliacijos modelius.

Kryžminė patikra vykdoma šiais etapais:

- Skaičiuojama visų išmatuotų taškų eksperimentinė variograma ir jos teorinis modelis.
- Kiekvienam taškui:
 - taškas pašalinamas iš duomenų rinkinio;
 - *tame taške* pagal *kitus taškus* ir sumodeliuotą variogramą skaičiuojama prognozė.
- Sumuojami modelio reikšmių nuokrypiai nuo tikrųjų reikšmių. Modelius galima lyginti pagal bendrąsias paklaidų statistikas, atsižvelgiant ir į svarbias atskiras prognozes. Kryžminei patikrai tinka šios paklaidų statistikos:
 - *vidutinę kvadratinę paklaidą* – kuo ji mažesnė, tuo geriau; skaičiuojama atliekant nepriklausomą patikrą;
 - *vidutinę standartinę prognozės paklaidą* – liekanų nuokrypis nuo krigingo dispersijos; skaičiuojama atliekant nepriklausomą patikrą.

Prognozėms ir tikrosioms reikšmėms palyginti skirti kitų tipų grafikai ir bendrosios statistikos ir patikros, ir kryžminės patikros atveju yra panašūs.



4.35 pav. Atliekant kryžminę patikrą, po vieną pašalinami visi išmatuoti taškai (raudonas taškas) ir sukuriami modeliai be šių taškų

OID	X	Y	Measured	Predicted	StdError	Error	Stdd_Error
0	383680.079101	6157932.69957	115	116.033054	12.567996	1.033054	0.082197
1	390703.152866	6182291.25532	112	103.786094	12.289347	-8.213906	-0.668376
2	398391.295217	6194429.1409	104	95.422069	12.539702	-8.577931	-0.684062
3	398784.398972	6160394.00831	110	106.853414	12.111396	-3.146586	-0.259804
4	403390.150312	6141106.71957	87	113.10013	12.319369	26.10013	2.118626
5	411908.839591	6154500.52427	118	93.727307	11.900164	-24.272693	-2.039694
6	421374.290348	6185861.6119	67	90.350601	12.105776	23.350601	1.928881
7	422300.95737	6137602.53501	90	96.342293	12.318489	6.342293	0.51486
8	429015.007483	6197588.50033	72	75.393714	12.514826	3.393714	0.271175
9	437502.470366	6172792.5159	76	75.766314	12.321599	-0.233686	-0.018966
10	446880.220164	6159746.69398	78	68.673814	12.566078	-9.326186	-0.742172

4.36 pav. Kryžminės patikros rezultatai: palygintos išmatuotos ir sumodeliuotos visų taškų reikšmės

Jei taškų pakankamai daug, pašalinus vieną tašką modelis neturėtų žymiai pasikeisti.

4.5.5 Modelių lyginimas

Kad būtų galima pasirinkti „geriausią“ paviršių, paprastai pirmiausia sukuriama keli paviršiai. Kiekvieną paviršių galima sistemingai palyginti su kitais pagal prognozės paklaidas. Geriausi iš lyginamų modelių bus tie, kurių *vidutinė standartinė prognozės paklaida* arčiausiai nulio, *vidutinė kvadratinė paklaida* mažiausia, *vidutinė prognozės paklaida* artimiausia *vidutinei kvadratinei paklaidai*, o *standartinė vidutinė kvadratinė prognozės paklaida* artimiausia vienetui.

Lyginant skirtingų metodų ir (arba) modelių rezultatus, reikia įvertinti du dalykus – optimalumą ir tinkamumą (arba teisingą kintamumą). Svarbu, kad gautas kintamumas būtų teisingas. Krigingo metodu prognozuojamos reikšmės priklauso nuo krigingo prognozės paklaidų $\hat{\sigma}_i$.

Pavyzdžiui, tam tikro modelio vidutinė kvadratinė paklaida gali būti mažesnė. Taigi, šį modelį lyg ir galėtume laikyti optimaliu. Bet palyginę jį su kitu modeliu nustatome, kad antrojo modelio vidutinė kvadratinė paklaida artimesnė vidutinei apskaičiuotai prognozės paklaidai. Jei *vidutinė krigingo prognozės paklaida* arčiau vidutinės kvadratinės prognozės paklaidos, šis modelis tinkamesnis, nes tik apskaičiuota krigingo prognozės paklaida įvertina nuo tikrųjų duomenų reikšmių nepriklausantį prognozės neapibrėžtumą (reikia tik variogramos modelio). Jei *vidutinė krigingo paklaida* didesnė už vidutinę kvadratinę

prognozės paklaidą, gautas prognozės kintamumas yra per didelis; jei *vidutinė krigingo paklaida* mažesnė už vidutinę kvadratinę prognozės paklaidą, gautas prognozės kintamumas yra per mažas.

Todėl prognozę kintamumą tinkamai įvertinti galima pagal *vidutinę kvadratinę standartinę prognozės paklaidą*. Taigi, jei *vidutinė krigingo paklaida* tinkamos, vidutinė kvadratinė standartinė prognozės paklaida bus artima 1.

4.5.6 Duomenų imties klausimai

Imties dydis ir skirstinys gali turėti įtakos prognozės tikslumui. Apskritai, kuo daugiau matuojamų taškų, tuo geriau – padidės imties tankis (matavimai bus arčiau vienas kito), bus galima geriau įvertinti vietinį kintamumą ir rezultatai bus teisingesni stambiu masteliu (mažose srityse). Tačiau padidės duomenų sukaupimo sąnaudos. Jei matavimų tankis mažas, sumažės jautrumas vietiniams pokyčiams ir bus įvertintas tik kintamumas visos srities mastu; rezultatai geriau tiks tyrimams smulkiu masteliu (didelių sričių tyrimams).

Imties formavimo tikslas – nustatyti erdvinės priklausomybės struktūrą (pvz., pusvariogramę) kuo mažesniu matavimų skaičiumi, t. y. gauti optimalų matuojamų taškų išdėstymą žemėlapyje. Imties sąnaudų minimizavimo sprendimai pateikti pačiuose krigingo metoduose.

Kaip jau minėjome, krigingo skaičiavimų paklaida priklauso *tik* nuo imties *konfigūracijos* ir pasirinkto erdvinės priklausomybės *modelio*, ir nepriklauso nuo duomenų reikšmių. Todėl, *jei* žinoma erdvinė struktūra (variogramos modelis), didžiausias arba vidutinės prognozės paklaidas galima nustatyti *prieš* skaičiuojant imtį. Remiantis prognozės paklaidomis ir *ekonominės naudos* kriterijais, imties formavimo sprendimus galima priimti prieš atliekant lauko tyrimus.

Optimalią taškų konfigūraciją galima sudaryti remiantis žinomų krigingo sistemų *optimizavimo kriterijais*, kuriuos galima laikyti tam tikrais kiekiniais sudarytos imties kokybės matais:

- Minimizuokite *didžiausią* krigingo *prognozės paklaidą* tiriamojoje srityje.
- Minimizuokite viso tiriamosios srities *vidutinę* krigingo *prognozės paklaidą*.
- Maksimizuokite *imties variogramos informacijos kiekį*, kad variogramą būtų įmanoma tinkamai įvertinti.

4.6 Geostatistinių metodų palyginimas

Šioje dalyje aptartos naujos sąvokos susijusios su *erdvine priklausomybe* arba *koreliacija erdvėje*, kuri bendru atveju reiškia, kad „kintamojo reikšmė viename erdvės ar laiko taške susijusi su jo reikšmėmis artimuose taškuose“. Žinant kintamojo reikšmes vienuose taškuose, galima numatyti (tam tikru pasikliautimumu) jo reikšmę bet kuriame kitame taške.

Šiuo atveju kintamųjų koreliacijos sąvoka taikoma *tam pačiam kintamajam*, o erdviniai ryšiai modeliuojami kaip *atstumo* arba *laiko* funkcijos. Ši koreliacija vadinama *autokoreliacija*. Pagrindinis klausimas buvo: kaip ją aprašyti? Erdvinio ryšio pobūdį atspindi erdvinė *struktūra*: ar toli, ir kuria kryptimi, pasireiškia erdvinė priklausomybė? Kaip ši priklausomybė keičiasi keičiantis atstumui ir kryptčiai tarp taškų?

Kiekybinė autokoreliacijos išraiška yra atstumo ir krypties tarp dviejų taškų funkcija. Erdvinė struktūra aprašoma aprėptimi, kryptimi ir stiprumu. Kokio tipo interpoliavimo metodą naudoti, priklauso nuo daugelio veiksnių. Paprastai išbandomi keli interpoliavimo metodai, ir palyginus rezultatus išsirenkamas geriausiai tinkantis konkrečiam atvejui. Tačiau interpoliavimo metodo pasirinkimas gali priklausyti ir nuo praktinių nagrinėjamos temos žinių.

Taip pat metodo pasirinkimas gali priklausyti ir nuo matuojamų taškų imties kokybės. Nuo imties atramos arerdvės srities, kurioje skaičiuojamos ar matuojamos vidutinės reikšmės, priklausys prognozės rezultatų skiriamoji geba.

Interpoliavimo metodų savybės apibendrintos 4.5 lentelėje.

4.5 lentelė. Kai kurios interpoliavimo metodų savybės

Metodas	Interpoliatoriaus tipas	Išvesties žemėlapiu tipas	Pranašumai	Trūkumai	Prielaidos
Atvirkščiai proporcingo atstumo	Determinuotasis interpoliatorius	Prognozės	Mažai skaičiuojamų parametrų	Neįvertinamos tikslumo paklaidos; aplink duomenų vietas atsiranda „dėmės“	Nereikia
Globali polinominė interpoliacija	Determinuotasis interpoliatorius	Prognozės	Mažai skaičiuojamų parametrų	Neįvertinamos matavimų paklaidos; rezultatas gali būti pernelyg glodus – didelės įtakos turi kraštiniai taškai	Nereikia
Lokali polinominė interpoliacija	Determinuotasis interpoliatorius	Prognozės	Daugiau skaičiuojamų parametrų	Neįvertinamos matavimų paklaidos; gali būti pernelyg automatinis	Nereikia
Splainai	Determinuotasis glodintuvas	Prognozės	Lankstus ir automatinis, skaičiuojami keli parametrai	Neįvertinamos matavimų paklaidos; gali būti pernelyg automatinis	Nereikia
Krigingas	Stochastinis prognozės įrankis	Prognozės, paklaidų, tikimybių, kvantilių	Labai lankstus: galima įvertinti autokoreliaciją erdvėje, apskaičiuoti prognozės paklaidas; skaičiuojama daug parametrų	Reikia priimti daug sprendimų – pasirinkti transformacijų tendus, modelius, parametrus ir kaimynystes	Kai kurie metodai tinka tik normaliai pasiskirsčiusiems duomenims
Kokrigingas	Stochastinis prognozės įrankis	Prognozės, paklaidų, tikimybių, kvantilių	Labai lankstus: galima apdoroti kelių duomenų rinkinių informaciją, įvertinti kryžminę koreliaciją erdvėje; skaičiuojama daug parametrų	Reikia priimti daug sprendimų – pasirinkti transformacijų tendus, modelius, parametrus ir kaimynystes	Kai kurie metodai tinka tik normaliai pasiskirsčiusiems duomenims

Dalies klausimai savarankiškam darbui:

- 1 Apibrėžkite, kas yra autokoreliacija.
- 2 Paašškinkite, kuo skiriasi globalūs ir vietiniai interpoliavimo metodai.
- 3 Kas yra tikslus interpoliavimo metodas?
- 4 Paašškinkite, kaip pusdispersė apibūdina erdvinę priklausomybę.
- 5 Apibrėžkite, kas yra pusvariogramės grynuolis, slenksčio atstumas ir slenkstis.
- 6 Kodėl pusvariogramei reikia pritaikyti matematinį modelį?
- 7 Kuo universalusis krigingas skiriasi nuo normaliojo?
- 8 Aprašykite, kaip atliekama kryžminė patikra.

Siūloma literatūra:

- Geostatistical Analyst, ESRI ArcGIS 9.2 Desktop Help,
[http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=An overview of Geostatistical Analyst](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=An+overview+of+Geostatistical+Analyst).
- Deterministic and Geostatistical Interpolation methods sections, Geospatial Analysis: Web site, M. J. de Smith, M. F. Goodchild, P. A. Longley,
<http://www.spatialanalysisonline.com/output/> .

Literatūra

1. Johnston K., Ver Hoef J. M., Krivoruchko K., and Lucas N. (2003) ArcGIS® 9 Using ArcGIS® Geostatistical Analyst, ESRI.
2. Berke, O. (1999) Estimation and Prediction in the Spatial Linear Model. Water, Air, and Soil Pollution 110, 215-237.
3. Kang-Tsung Chang (2006) Introduction to Geographic Information Systems, Third Edition, The McGraw Hill.
4. Bailey, T. & Gatrell, A. (1995) Interactive Spatial Data Analysis. Addison Wesley Longman, Harlow.
5. Cressie, N. (1993) Statistics for Spatial Data (Revised Edition). Wiley, New York.

Vartojami terminai

<i>Geostatistics</i>	– geostatistika
<i>Regionalized variables</i>	– regioniniai kintamieji
<i>Auto-correlation</i>	– autokoreliacija
<i>Spatial dependence</i>	– erdvinė priklausomybė
<i>Inverse distance weighting</i>	– atvirkščiai proporcingo atstumo metodas
<i>Trend</i>	– trendas
<i>Global and local polynomial</i>	– globali ir vietinė polinominė interpoliacija
<i>Splines</i>	– splainai
<i>Kriging</i>	– krigingas
<i>Regression</i>	– regresija
<i>Power function</i>	– laipsninė funkcija
<i>Neighborhood search strategy</i>	– paieškos kaimynystėje strategija
<i>Omnidirectional and isotropy</i>	– įvairiakryptiškumas ir izotropija
<i>Anisotropy</i>	– anizotropija
<i>Semivariance</i>	– pusdispersė
<i>Variogram</i>	– variograma
<i>Empirical semivariogram</i>	– empirinė pusvariogramė
<i>Lag bins</i>	– atskyrimo grupės
<i>Auto-covariance</i>	– autokovariacija
<i>Theoretical, authorized or fitted variogram</i>	– teorinė, leistina arba pritaikyta variograma
<i>Ordinary, simple, universal, regression kriging and co-kriging</i>	– normalusis, paprastasis, universalusis, regresinis krigingas ir kokrigingas
<i>Unbiased</i>	– be sisteminės paklaidos
<i>Validation</i>	– patikra
<i>Cross-validation</i>	– kryžminė patikra
<i>RMSE</i>	– VKP (vidutinė kvadratinė paklaida)
<i>ME</i>	– VN
<i>Kriging prediction error</i>	– krigingo prognozės paklaida
<i>Average kriging prediction error</i>	– vidutinė krigingo prognozės paklaida
<i>Root-mean-square standardized prediction error</i>	– vidutinė kvadratinė standartinė prognozės paklaida
<i>Error, quantile and probability maps</i>	– paklaidų, kvantilių ir tikimybių žemėlapiai

5. Erdvinis modeliavimas

Paprasčiausias geografinės informacinės sistemos (GIS) taikymo atvejis – naudoti ją duomenų rinkiniams organizuoti ir redaguoti bei rezultatams parodyti skaitmeninio formato žemėlapyje, kurį galima peržiūrėti arba atspausdinti. Tačiau turbūt svarbiausia GIS savybė yra sugebėjimas kurti modelius tikriems įvykiams imituoti. Nors kiekvienas žemėlapis ir taip yra erdvinis modelis, nes jame atvaizduota tikrovė, bet GIS galima išanalizuoti kelis kintamuosius ir interpretuojant šios analizės rezultatus imituoti arba numatyti galimus įvykius. Ši galimybė leidžia interpoliuoti informaciją didelėse teritorijose. Pavyzdžiui, būtų nepraktiška (t. y. per brangu) atlikti visos šalies gyvūnijos buveinių lauko tyrimus, bet GIS galima sumodeliuoti šių buveinių tinkamumą dideliems plotams.

Šioje kurso dalyje nagrinėjamos kai kurios pagrindinės erdvinio modeliavimo funkcijos.

- Erdvinių problemų modeliavimas
- Erdvinių modelių klasifikavimas
- Modelių tipai
- Modelių kūrimo priemonės
- Modelių pavyzdžiai

5.1 Erdvinių problemų modeliavimas

Erdvinis modeliavimas yra sudėtingiausias erdvinės analizės lygmuo. Modeliai leidžia supaprastintai atvaizduoti kai kuriuos tikrovės aspektus. Kaip abstraktūs arba daliniai atvaizdžiai modeliai padeda paaiškinti arba numatyti, kaip tikrovėje vystosi (arba gali vystytis) tam tikra žmogaus veikla, natūralus reiškinyje arba sistema. Jais galima vaizduoti buvusias, esamas arba būsimas sąlygas. Paprastai modelį sudaro tam tikri erdviniai kintamieji, susieti tarpusavyje aritmetine operacija arba tam tikromis komandomis.

Modelio kūrimą galima suskirstyti į kelis etapus – nuo uždavinio formulavimo iki rezultatų taikymo. Kuriant modelį ypač svarbu suprasti uždavinį konceptuali lygmeniu, todėl prieš pradedant dirbti kompiuteriu primygtinai rekomenduojame atlikti toliau išvardytus veiksmus.

1 etapas. Uždavinio formulavimas: norint sukurti konkrečiai užduočiai spręsti skirtą modelį, reikia aiškiai apibrėžti pačią užduotį ir modelio tikslą (-us). Šiuo etapu nustatomi modelio parametrai. Jie nustatomi atsakant į šiuos klausimus:

- Kokie reiškiniai modeliuojami?
- Kodėl reikia kurti modelį?
- Koks modelio mastas ir aprėptis (t. y. ar nustatyta tyrimų sritis)?
- Koks modelio laiko intervalas – ar įvertinamas vienas, ar keli periodai?

Pavyzdžiui, vertindamas poveikį aplinkai, biologas gali norėti nustatyti, kokiam buveinių plotui planuojamo projekto įtakos zonoje gali būti padaryta žalos. Šiam uždaviniui spręsti galima naudoti modelį (galbūt kelis modelius), apjungiantį su gyvūnų ir augalų buveinėmis susijusius kintamuosius. Šiuo atveju pirmasis uždavinio formulavimo veiksmas būtų nustatyti, kokių rūšių gali būti projekto srityje ir išsirinkti rūšis, kurioms bus kuriami buveinių modeliai. Pavyzdžiui, gali būti nustatyta, kad įvairioms esamos gyvūnijos rūšims modeliuoti užteks šešių kontrolinių rūšių (pvz., plėšrūno, aukos, paukščio, roplio, kailinio gyvūno ir kanopinio). Šios rūšys ir taps modeliuojamais reiškiniais. Modeliuojant gyvūnijos buveines, tiriamos srities dydis paprastai nustatomas pagal didžiausią rūšies individualios teritorijos spindulį arba siūlomo projekto sritį apimančius vandens baseinus ar ekosistemas. Šie ribų tipai gali padėti apibrėžti modelio ribas erdvėje. Paprastas būdas apskaičiuoti galinčių išnykti buveinių plotą – įvertinti buveines skirtingais periodais. Šiuo atveju tai būtų atlikta sumodeliavus esamas (pradines) sąlygas ir palyginus jas su antru modelio rezultatų rinkiniu, atitinkančiu buveinės įgyvendinimą projektą.

2 etapas. Uždavinio išskaidymas: nustatę uždavinį ir tikslus, uždavinį galima išskaidyti į sudedamąsias dalis, kad modeliavimo veiksmus būtų patogiau valdyti. Tam reikia nustatyti tikslą pasiekti padėsiančias užduotis, susijusias reiškiniais ir šių reiškinų tarpusavio sąveiką. Reiškinių nustatymas leidžia modeliotojui nustatyti, kokių duomenų rinkinių reikės, ir surinkti šiuos duomenis. Kad erdviniai ir atributiniai ryšiai tarp sudedamųjų modelio dalių taptų aiškesni, dažnai naudinga juos vaizduoti blokine schema.

Nagrinėjamu gyvūnų buveinių įvertinimo atveju modeliotojas įvertintų kiekvienos modeliuojamos rūšies buveinės reikalavimus (pvz., aukštį virš jūros lygio, nuolydį, vandens artumą, augmeniją) ir sudarytų atitinkamų duomenų rinkinių katalogą bei apibrėžtų, kaip nuo šių duomenų atributų ir erdvinių ryšių priklauso galimos kiekvienos rūšies buveinės. Pavyzdžiui, jei žinoma, kad rūšis gyvena ne aukščiau negu 500 metrų, ant ne statesnio

negu 10 proc. paviršiaus, 250 metrų spinduliu nuo gėlo vandens šaltinio ir mėgsta spygliuočių mišką, iš skaitmeninio aukščių modelio (DEM) galima nustatyti aukščio ir nuolydžio reikalavimus tenkinančias sritis. Kad galėtume modeliuoti vandens artumą, reikės hidrologijos sluoksnio, o buveinėms spygliuočių miškuose nustatyti tiks augmenijos arba žemės dangos duomenų rinkinys. Šių keturių sluoksnių sankirta leis nustatyti ir galimas dominančių rūšių buveines.

3 etapas. Modelio kūrimas ir kalibravimas: šiuo etapu nustatomos priemonės ir matematinės operacijos, kurių reikės analizei atlikti. Naudodamas 2 etape išnagrinėtus duomenis, modeliotojas iš šių priemonių ir operacijų sukuria modelį. Tada šį modelį galima kelis kartus įvykdyti ir sukalibruoti. Kalibruojant modelio rezultatai lyginami su įvesties duomenimis, ir matematinių operacijų parametrai derinami taip, kad rezultatai būtų kuo tikslesni.

Dirbdamas su GIS, modeliotojas įvykdytų modelį kelis kartus ir gautus rezultatus palygintų su duomenimis. Modelio rezultatai tikslinami derinant parametrus ir operacijas.

Nagrinėjamu gyvūnų buveinių žemėlapiu atveju pradiniai rezultatai vaizduotų galimas buveines, apskaičiuotas pagal pasirinktus kintamuosius; tada šiuos rezultatus galima tikslinti pagal kitus kriterijus. Pavyzdžiui, galbūt ideali buveinė (gerai tinkama) yra 100 metrų atstumu nuo gėlo vandens šaltinio, o vidutiniškai tinkama – nuo 100 iki 250 metrų atstumu.

4 etapas. Modelio rezultatų patikrinimas: įvertinimas, ar modelis gali tiksliai numatyti tikrovės reiškinius. Tikrinant modelį, jo rezultatai lyginami su lauko tyrimų duomenimis ir (arba) vykdant modelį su kitais duomenimis, nepanašiais į naudotus kalibruojant. Jei kito duomenų rinkinio nėra, galima padalinti vieną duomenų rinkinį į dvi dalis, ir vieną naudoti modeliui kurti ir kalibruoti, o kitą – modeliui tikrinti.

Klaidų nustatymas

Apžiūra. Vienas iš būdų nustatyti klaidas – tiesiog pažvelgti į rezultatus ir įvertinti, ar jie atrodo logiški, pastovūs ir ar atitinka įvestus duomenis.

Dokumentavimas. Ar duomenys naudojami tinkamai, galima nustatyti pagal įvesties duomenų rinkinių metaduomenis – mastelio, tikslumo, atributų reikšmes.

Tikrinimo taisyklės. Dvi naudingos taisyklės, pagal kurias galima patikrinti duomenų tinkamumą modeliui, yra topologinių ir atributų domenų intervalai.

Rezultatų pastovumas. Galima iš naujo surinkti duomenis, pakartoti konvertavimo bei apdorojimo veiksmus ir palyginti modelio rezultatus.

Patikra vietoje. Patikimas modelio tikrinimo būdas – patikrinti jo rezultatus lauko tyrimais.

Statistika. Statistinis tikrinimo būdas – skaičiuoti modelio rezultatų koreliaciją su stipriai susijusiu kintamuoju.

Nagrinėjamu atveju biologas galėtų patikrinti buveinių tinkamumo modelio rezultatus lygindamas gautas didelio tinkamumo buveinių sritis su žinomomis gyvūnų buveinių vietomis (pvz., guoliavietėmis).

5 etapas. Modelio rezultatų pritaikymas: patikrinęs modelį, modeliotojas gali naudotis modelio rezultatais.

Pagal patikrintus modelio rezultatus jis gali nustatyti dabartines buveines, tada perdengti jas siūlomo projekto įtakos sritimi ir apskaičiuoti, kiek buveinių bus prarasta įgyvendinus siūlomą projektą. Be to, pagal modelio rezultatus galima būtų numatyti prioritetines būsimų lauko tyrimų sritis.

Modeliavimo trūkumai

Labai svarbu suvokti, kad kiekvienas modelis turi trūkumų. Veiksniai, galintys apriboti modelio taikymą:

- nepakanka duomenų – pavyzdžiui, kai kurie arba visi tiriamosios srities duomenys gali būti nepakankamai išsamūs;
- naudotojas neteisingai supranta duomenis – pavyzdžiui, daug modelių nevaizduoja tiesiogiai interpretuojamų reikšmių, o yra jų tikimybės analizė;
- netinkamas modelis – šiuo atveju modelio rezultatai neteisingi dėl klaidingų prielaidų kuriant modelį arba dėl klaidų jį vykdant.

5.2 Erdvinių modelių klasifikavimas

Modelius galima klasifikuoti pagal paskirtį, metodiką arba logiką, o kai kurios klasifikacijos apima kelias kategorijas.

5.2.1 Paskirtis

Aprašomasis modelis

Aprašomasis modelis aprašo dabartinės tikro pasaulio (gamtinės arba antropogeninės aplinkos) sąlygas. Aprašomuoju modeliu galima pavadinti ir paprastą teminį žemėnaudos žemėlapi, nes jis vaizduoja tam tikrą žemės paviršiaus charakteristiką. Kiti pavyzdžiai gali būti skaitmeninis aukščių modelis, žemės danga, meteorologiniai žemėlapiai ir augmenijos žemėlapiai.

Aiškinamasis modelis

Kaip galima spręsti iš pavadinimo, aiškinamieji modeliai bando paaiškinti arba vykstančius reiškinius, nustatydami įtakos turinčius veiksniai ir įvertindami jų santykinę svarbą. Aiškinamųjų modelių pavyzdžiai yra dirvos erozijos, ozono sluoksnio suplonėjimo ir dumblių žydėjimo modeliai.

Prognozuojamasis modeliai

Prognozuojamasis modelis numato (pagal aiškinamuoju modeliu nustatytus veiksniai), kur gali vykti tam tikri aplinkos reiškiniai. Pavyzdžiui, iš augmenijos ir nuolydžių rastrų gautu buveinių tinkamumo žemėlapiu galima remtis tiriant, ar sritis tiks tam tikrų gyvūnų buveinėms. Kitas prognozuojamojo modelio pavyzdys būtų jūros lygio kilimo numatymas pagal klimato (pvz., vidutinių temperatūros reikšmių) pokyčius.

Norminamasis modelis

Norminamieji modeliai bando nurodyti, kaip turi vystytis tikrovės reiškiniai (ypač tinklai); jie siūlo optimalius sprendimus konkrečiomis sąlygomis. Šių modelių pavyzdžiai yra maisto pagalbos paskirstymas, eismo intensyvumas, maršrutų planavimas (pvz., keliaujančių pardavėjų uždavinys).

5.2.2 Metodika

Stochastiniai

Stochastinis (tikimybinis) modelis vaizduojamas matematine lygtimi, kurios bent vienas parametras tam tikru laipsniu yra atsitiktinis. Dėl šio atsitiktinumo atsiranda tam tikra modelio rezultatų paklaida arba neapibrėžtumas, kuris laikomas neišvengiamu ir išreiškiamas tikimybe. Pavyzdys būtų modelis, įvertinantis nuošliaužų tikimybę iškirtus tam tikras miško sritis. Kitas stochastinio modelio pavyzdys būtų krigingo analizės taikymas vandens mėginių matavimams. Gautume interpoliuotą paviršių, vaizduojantį galimą tam tikro teršalo koncentraciją (prognozuojamą žemėlapi) bei standartines visų numatomų reikšmių paklaidas.

Deterministiniai

Priešingai negu stochastiniai, deterministiniai modeliai laiko, kad nė vienas iš matematinės lygties kintamųjų arba parametru nėra atsitiktinis.

5.2.3 Statiniai arba dinaminiai

Statinis modelis nagrinėja reiškinius vienu laiko momentu, o dinaminis nagrinėja ir išryškina reiškinių pokyčius laikui bėgant bei sąveiką tarp kintamųjų. Pavyzdžiui, statinis modelis gali sumuoti analizuojamuose vienetuose esančių kelių tankį kuriais nors metais. Dinaminis modelis pagal šio modelio rezultatus nustatytų metinius kelių tankio pokyčius.

5.2.4 Pagal logiką

Indukciniai

Indukciniai modeliai yra perėjimas nuo konkretaus prie bendro, jų išvados grindžiamos ankstesnių tyrimų rezultatais. Paprastai indukciniai modeliai bando nustatyti bendrus dėsningumus arba taisykles tuo atveju, kai nėra visiškai suprantamos svarbios temos arba ryšiai. Indukcinio mąstymo pavyzdys yra vien tik dabartinėmis radinių vietomis pagrįstas archeologinio potencialo modelis.

Dedukciniai

Dedukciniai modeliai daro konkrečias išvadas iš bendrų prielaidų – mokslinių teorijų arba fizikos dėsnių, kai kintamieji ir jų sąveika gerai žinomi. Dedukcinio mąstymo pavyzdys yra archeologinio potencialo modelis, numatantis kasinėjimų vietas pagal fizinius apribojimus (pvz., nuolydį, padėtį, vandens artumą).

5.3 Modelių tipai

5.3.1 Vaizduojamieji modeliai

Dvejetainiai

Dvejetainiai modeliai pagal logines išraiškas atrenka sluoksnių arba rastrinių duomenų rinkinių elementus – panašiai kaip užklauso. Rezultatas yra rastras arba vektorinis sluoksnis, turintis dvi reikšmes: kriterijus atitinkantiems erdviniams elementams arba rastro gardelėms suteikiama reikšmė 1 (teisinga), kriterijų neatitinkantiems – 0 (klaidinga). Dvejetainiais modeliais atliekama analizė gali būti ir rastrinė, ir vektorinė.

Vektorinis dvejetainis modelis įvesties duomenų elementus ir jų atributus kombinuoja perdangos operacijomis (pvz., sankirta, sąjunga).

Rastrinis dvejetainis modelis atrenka duomenis pagal norimus kriterijus tiesioginėmis kelių rastrinių sluoksnių užklausomis.

Labiausiai paplitęs dvejetainio rastrinio modelio taikymas yra vietos išrinkimas, kai įvertinami keli kriterijai ir nustatoma tinkamiausia tam tikram objektui arba tam tikrai žemėnaudai vieta. Vietos išrinkimo analizė dažniausiai atliekama vienu iš dviejų būdų: arba pagal griežtus kriterijus vertinamos kelios iš anksto pasirinktos vietos, arba vertinamos visos įmanomos vietos.

Tarkime, kad kasybos įmonė slėnyje nori atrinkti galimas aukso kasyklų vietas pagal šiuos kriterijus:

- tam tikra mažiausia rūdos koncentracija;
- mažiausias atviros kasyklos dydis – 10 hektarų;
- mažiausia užliejimo tikimybė;
- kasyklos turi būti ne toliau negu 30 kilometrų nuo esamo kelio;
- kasyklos teritorijos paviršiaus nuolydis neturi būti didesnis negu 10 procentų.

Modelio kūrimo etapai:

- surinkti ir paruošti analizei visus reikiamus duomenų rinkinius (pvz., iš skaitmeninio aukščių modelio gauti dvejetainį rastrą, vaizduojantį ne didesnio negu 10 proc. nuolydžio sritis, ir konvertuoti šį rastrą į vektorinį poligonų sluoksnį);
- sukurti 30 kilometrų buferio zoną aplink visus kelius;
- perdangos funkcija (pvz., sąjungos) sujungti su gautais kelių buferiais kitus sluoksnius;
- atlikti gauto jungtinio sluoksnio užklausą ir nustatyti, kurios vietos atitinka anksčiau išvardytus kriterijus.

Ranginiai

Dar vadinami indeksų modeliais, ranginiai modeliai skaičiuoja kiekvieno nagrinėjamo ploto vieneto indekso vertę ir sukuria ranginį žemėlapi. Kaip ir dvejetainiuose modeliuose, ranginiuose modeliuose atliekamos perdangos bei rastrų aritmetikos operacijos ir įvertinami keli kriterijai. Skirtumas tas, kad dvejetainių modelių rezultatas yra „taip“ arba „ne“ (1 arba 0), ranginių modelių rezultatas yra nagrinėjamų vienetų indekso reikšmės.

Paprastai indeksų vertės skaičiuojamos svorinių tiesinių kombinacijų (*weighted linear combination*) metodu. Šis metodas atliekamas trimis etapais:

- nustatoma santykinė svarba – kriterijai lyginami tarpusavyje, ir jų santykinė svarba išreiškiama svoriais nuo 0,0 iki 1,0 (0–100 proc.);
- standartizuojami kriterijai – standartizuojamos santykinės kriterijų svarbos reikšmės; šis etapas palengvina kelių kintamųjų lyginimą;
- skaičiuojamos indekso reikšmės – nagrinėjamo ploto vieneto indekso vertė gaunama sujungus visas svertines kriterijų reikšmes ir padalinus gautą skaičių iš visos svorių sumos.

Ranginiai modeliai daugiausia naudojami kuriant tinkamumo arba pažeidžiamumo žemėlapius, kuriuos aptarsime nagrinėdami modelių pavyzdžius.

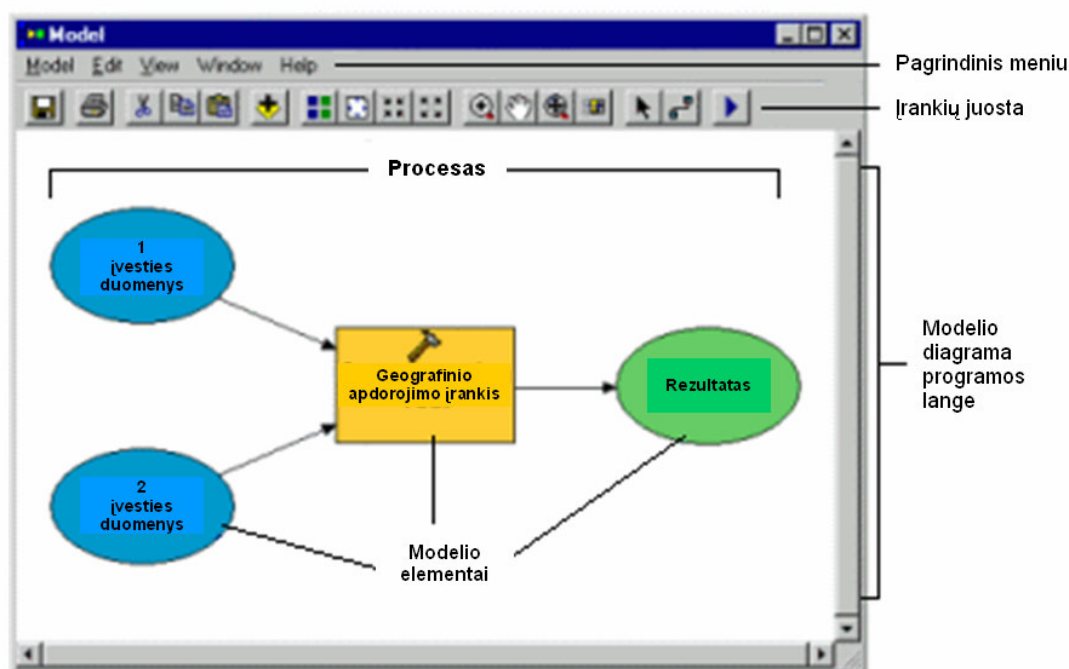
5.3.2 Procesų modeliai

Procesų modeliai imituoja tikrovėje vykstančius procesus, apjungdami turimą erdvinę informaciją. Kadangi šie modeliai skirti interpretuoti kelių kintamųjų sąveiką ir numatyti galimus įvykius, jie klasifikuojami ir kaip dinaminiai, ir kaip prognozuojamieji. Paprastai procesų modeliais analizė atliekama rastrinėje GIS aplinkoje. Rastrinėje aplinkoje galima atlikti sudėtingesnius aritmetinius skaičiavimus bei standartizuoti nesugretinamus duomenų šaltinius. Geri proceso modelių pavyzdžiai yra 5.5.4 skyriuje išsamiai nagrinėjami dirvos erozijos modeliai.

5.4 Modelių kūrimo programa Model Builder

Modelių kūrimo programa *Model Builder* – tai ESRI *ArcGIS* sąsaja, skirta sujungti kelis procesus ir palengvinti modelių kūrimą. Ji leidžia vizualizuoti darbo eigą (blokinėmis schemomis) bei kurti ir automatizuoti geografinio duomenų apdorojimo užduotis, kurias kitu atveju tektų atlikti *ArcMap* programa rankiniu būdu. Iš to išplaukia dar vienas šios sąsajos pranašumas – galimybė dokumentuoti modelio kūrimo veiksmus. Nors šia sąsaja pradinį modelio variantą sukurti gali trukti kiek ilgiau negu atlikti veiksmus rankiniu būdu, ji ypač naudinga vykdant modelį kelis kartus – modelį galima vykdyti su skirtingais duomenimis arba pakoreguoti ir įvykdyti iš naujo, o tai leidžia nagrinėti modelio variantus bei prielaidas. *Model Builder* procesas gali būti arba *ArcGIS* sistemos įrankis (*ArcTools* įrankis), arba paties modeliautojo sukurtas įrankis.

Model Builder sąsaja – tai programos langas, kuriame yra modelio diagrama (proceso ir modelio elementai), pagrindinė meniu juosta su funkcijomis, o taip pat įrankių juosta su modelio elementams programos lange valdyti skirtais įrankiais bei funkcijomis (5.1 pav.).



5.1 pav. Model Builder sąsaja

Modelio elementai gali būti įvesties duomenys, geografinio apdorojimo įrankiai ir išvesties duomenų sluoksniai. Procesai kuriami velkant įrankius ir duomenis pele į programos langą ir įvedant įrankių parametrus. Sujungus įvesties duomenų kintamąjį su geografinio apdorojimo įrankiu, įvesties duomenų reikšmė (pvz. duomenų sluoksnio pavadinimas) automatiškai įrašoma į įrankio dialogo langą kaip įvesties parametras. Procesai atliekami paleidus vykdyti programos lange sudarytą modelį – nurodyti sluoksniai apdorojami nurodyta tvarka nurodytos konfigūracijos įrankiais.

Į programos langą įkėlus ir tinkamai sujungus norimus duomenis bei įrankius, modelį galima išsaugoti. Išsaugotą modelio konfigūraciją galima naudoti kaip šabloną ir keisti modelio parametrus ir (arba) įvedamus duomenis.

Minėtojo aukso kasyklų pasirinkimo pavyzdžio veiksmus galima įrašyti į *Model Builder* ir įvykdyti kaip užbaigtą automatizuotą procesą. *Model builder* lange galima sukurti nuolydžio sluoksnio ir kelių buferio procesus ir sujungti juos su sąjungos įrankiu – bus gautas sluoksnis, vaizduojantis tinkamas kasyklų vietas.

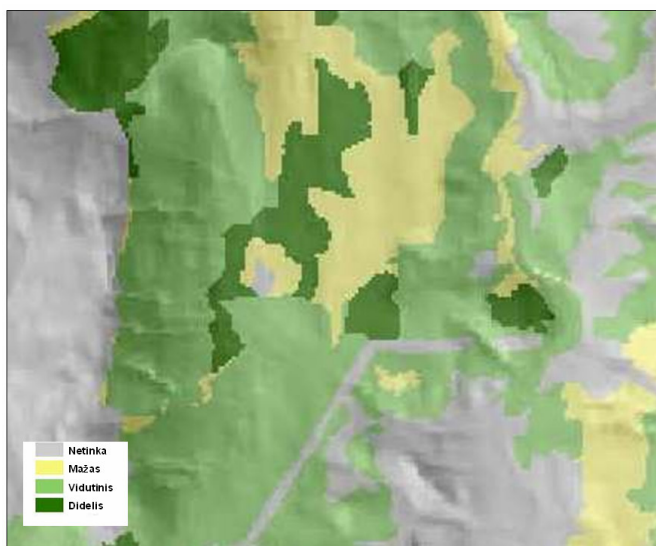
Kad *Model Builder* kuriami modeliai duotų norimą rezultatą, ypač svarbu po kiekvieno veiksmo pakartotinai patikrinti su kiekvienu modelio elementu susijusius apdorojimo veiksmus bei peržiūrėti išvedamus duomenis. Kaip ir kiekvieno automatizuoto proceso, modelio klaidas nustatyti gali būti sudėtinga, todėl šiuo atveju labai svarbios yra kokybės užtikrinimo ir kokybės kontrolės procedūros.

5.5 Modelių pavyzdžiai

5.5.1 Laukinių gyvūnų buveinės

Buveinių tinkamumas

Nustačius ankstesnę, dabartinę ir būsimą nagrinėjamos srities kintamųjų, ryšių bei priklausomybių būseną, galima įvertinti ekosistemos būklę bei funkcijas. Augmenijos ir gyvūnijos žemėlapiai yra naudinga planavimo bei valdymo priemonė, nes jie ne tik vaizduoja visos srities būklę, bet ir pateikia informaciją apie atskiras rūšis. Jie tinka ir kaip regioninė perspektyva (pvz., apribojimų žemėlapių kintamasis), ir konkrečiam tikslui (pvz., padėti pasirinkti mėginių ėmimo vietas lauko tyrimams). Žemėlapiuose vaizduojama buveinių vieta ir aprėptis bei nurodomas santykinis jų tinkamumas – priskiriamas buveinės tinkamumo indeksas arba klasė (5.2 pav.). Galima apjungti paviršiaus, aukščio, augmenijos, vandens objektų bei žmogaus keliamų trikdžių duomenis ir pavaizduoti buveinių tinkamumą atsižvelgiant į konkrečios buveinės reikalavimus. Pavyzdžiui, tam tikra rūšis gali labiau mėgti tam tikrą aukščių virš jūros lygio intervalą, augmenijos tipą ir gyventi ne toliau negu 200 metrų atstumu nuo gėlo vandens telkinio. Dirbant GIS, galima kiekybiškai įvertinti visus šiuos kintamuosius, sukurti modelį ir pritaikyti jį turimiems duomenims.



5.2 pav. Buveinių tinkamumo žemėlapių pavyzdys

Šiai analizei reikalingi duomenys yra:

- Ekologinės žemės klasifikacijos (ELC) žemėlapiai ir (arba) augmenijos duomenys
- Reljefas
 - nuolydis
 - ekspozicija
 - aukštis virš jūros lygio
- Bazinio žemėlapių informacija
 - hidrologija
 - žmonių keliami trikdžiai
- Dalykinės žinios
 - gyvūnijos ir augmenijos ekspertai

- GIS ir duomenų integravimas

Buveinių fragmentacija

Kita pagalbinė priemonė aplinkos pokyčiams valdyti yra buveinių fragmentacijos žemėlapiai. Buveinių fragmentacijos žemėlapiuose (*habitat fragmentation maps*) vaizduojami maži buveinių ploteliai bei vientisi buveinių plotai, gauti sujungus laukinių gyvūnų bei augalų buveinių sluoksnius su žmogaus keliamų trikdžių žemėlapiais. Šie jungtiniai žemėlapiai padeda apskaičiuoti perspektyvių buveinių plotą ir įvertinti galimus nuostolius dėl esamos ar siūlomos plėtros. Jie padeda nustatyti: visą buveinių plotą ir vidutinį buveinių fragmento plotą; galimą buveinių pakraščių ploto padidėjimą (tai teigiamas arba neigiamas daugiklis, priklausantis nuo nagrinėjamos rūšies); galimą vidinės buveinių dalies ploto sumažėjimą; izoliuotais galinčius tapti buveinių fragmentus; galimą mažesnių buveinių fragmentų ploto padidėjimą.

Šie rezultatai gali padėti nustatyti, kur yra pagrindinės kertinės buveinės ir buveinių salelės, nuo kurių priklauso rūšies išlikimas. Kertinė buveinė – tai pakankamai didelė (dydis priklauso nuo nagrinėjamų rūšių) ir ekologiškai nepažeista buveinės sritis, kurioje gali gyventi tam tikros rūšys, įskaitant ir tas, kurioms reikia vidinių buveinių. Buveinės salelė yra mažas buveinės fragmentas, teikiantis gyvūnams maistą, prieglobstį ar tenkinantis kitus poreikius. Tai gali būti ir žmogaus pakeista ekosistema. Salelės paprastai per mažos genetinei populiacijos įvairovei išlaikyti, bet tai gali būti savotiški „laipteliai“ į kertines buveines, svarbūs rūšims ir ekologinėms funkcijoms.

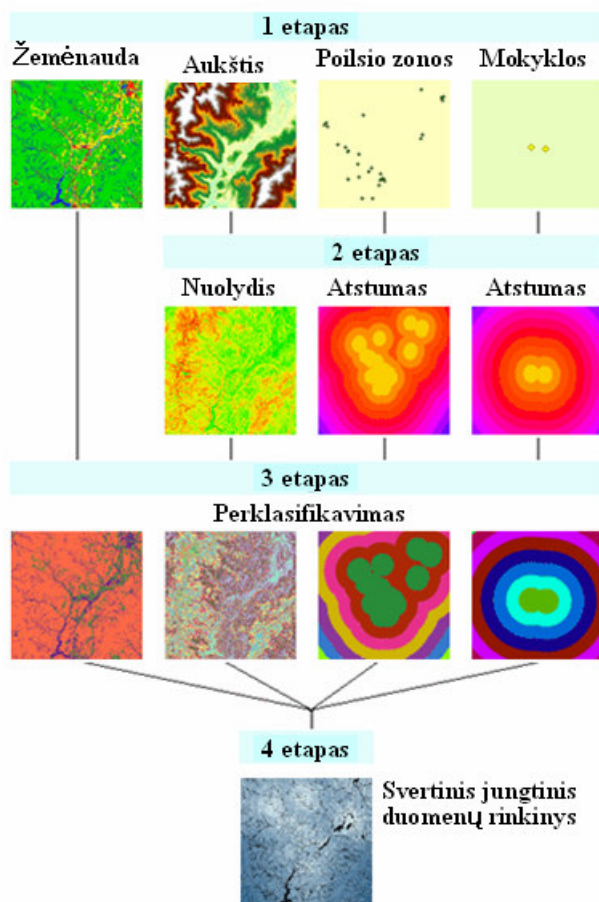
Šiai analizei reikalingi duomenys yra:

1. Ekologinės žemės klasifikacijos žemėlapiai ir (ar) augmenijos duomenys
2. Laukinių gyvūnų buveinių žemėlapiai
3. Reljefas (nebūtinai)
 1. nuolydis
 2. ekspozicija
 3. aukštis
4. Bazinio žemėlapių informacija
 1. žmonių keliama trikdžiai
5. Dalykinės žinios
 1. gyvūnijos ir augmenijos ekspertai
 2. GIS ir duomenų integravimas

5.5.2 Vietovės tinkamumo modeliai

Tinkamumo žemėlapiai

Kartais rastrai perklasifikuojami (keičiamos gardelių reikšmės) taip, kad gardelėms priskiriamos reikšmės arba svoriai atspindėtų svarbą, jautrumą arba hierarchiją. Šis procesas dažnai naudojamas kuriant tinkamumo žemėlapius (pvz., buveinių tinkamumo modelis iš kelių rastrinių įvesties sluoksnių, kuriuo gaunamos reikšmės nuo 1 (mažas tinkamumas) iki 6 (didelis tinkamumas) (5.3 pav.).



1 etapas. Apibrėžkite įvesties duomenis – nustatykite, kokie duomenys bus modelio įvesties sluoksniai.

2 etapas. Sukurkite išvestinius duomenų rinkinius – iš pradinių duomenų gaukite reikiamus rastrus (pvz., nuolydžio ir krypties duomenis galima išvesti iš aukščių rastro).

3 etapas. Perklasifikuokite – suvienodinkite įvesties duomenų rinkinių verčių intervalą, kad duomenis būtų galima lyginti ir jungti toje pačioje skalėje.

4 etapas. Priskirkite svorius ir sujunkite rastrus – suteikę didesnį svorį svarbesniems duomenų rinkiniams, sujunkite duomenų rinkinius, ir gausite tinkamumo rastrą.

5.3 pav. Perklasifikavimas – tinkamumo žemėlapių kūrimas

Perklasifikuojant visada sukuriamas naujas rastras, kas matyti ir iš pateikto pavyzdžio. Rastrai dažniausiai perklasifikuojami ruošiant duomenis jungti (atliekant perdangos analizę), kad sumažėtų išvesties kategorijų.

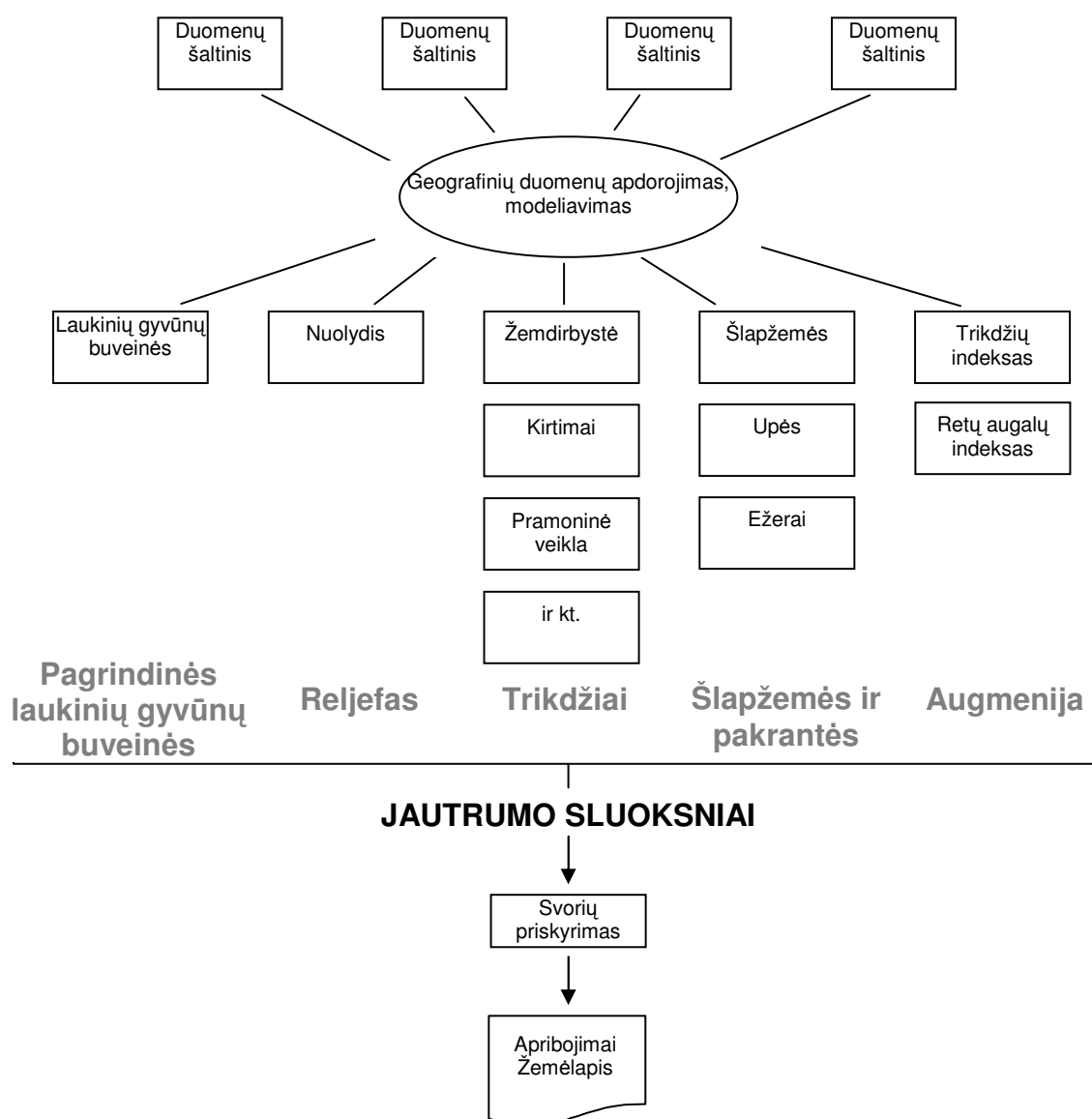
Apribojimų žemėlapiai

Apribojimų žemėlapiai yra vienas iš tinkamumo žemėlapių tipų. Ši analizė nustato būsimų projektų įgyvendinimo galimybes ir apribojimus. Surenkami įvairūs erdviniai duomenų sluoksniai (pvz., reljefo ir topografijos, augaviečių ir gyvūnų buveinių, nuolydžio, apsauginių zonų, dirvožemio, paveldo informacija), kiekvieno sluoksnio duomenims priskiriama jautrumo reikšmė (pagal mokslines ir vietines žinias) ir sujungiant sluoksnius gaunamas vienas žemėlapis, apjungiantis visus įvesties sluoksnius. Gautas išvestinis kartografinis produktas nustato ir suformuoja sudėtingus ryšius tarp skirtingų aplinkos duomenų rinkinių, ir įvertina vietos bei funkcinius apribojimus, kuriuos lemia projekto ypatybės. Apribojimų žemėlapiai yra vaizdinė priemonė, padedanti priimti sprendimus – juose nurodomos aplinkosaugos ir kultūros požiūriu svarbios sritys, atsižvelgiant į jautrių kraštovaizdžio objektų vietas. Apribojimai gali būti išreikšti kaip nurodymai iš viso vengti tam tikrų vietų (pvz., netinkama sritis) arba kaip laipsniški svarbos lygmenys. Idealiu atveju apribojimų žemėlapis padeda išrinkti projektui tokią vietą, kurioje apribojimų mažiausiai, ir taip sumažina pramoninės veiklos žalą aplinkai ir kultūrai.

Gaunamas išvestinis paviršius priklauso nuo įvesties duomenų reikšmių ir duomenims priskirtų svorių. Pavyzdžiui, nustatant, kur šiuo metu yra tinkamiausi gyvūnų migracijos koridoriai, gautas rezultatas priklausys nuo nuolydžio ir ekologinės žemės klasifikacijos (ELC) tipo, o prieinamumo žemėlapis priklausys nuo nuolydžio ir atstumų iki bendruomenių.

Apžvalga

Apribojimų žemėlapiai sujungia kelis erdvinius sluoksnius į vieną žemėlapi, nurodantį projekto vietos išrinkimo galimybes ir apribojimus. 5.4 pav. pateikta apribojimų žemėlapio sudarymo proceso apžvalga.



5.4 pav. Jautrumo sluoksnio struktūros pavyzdys

Apribojimų žemėlapiai sudaromi keliais etapais, o šioje kurso dalyje aptartos sąvokos tiesiogiai susijusios su toliau aprašoma ir naudojama modeliavimo metodika.

- Apribojimų žemėlapių yra vaizdinis apribojimų modelio rezultatas. Jie kuriami matematiškai sujungiant neribotą skaičių jautrumo sluoksnių.
- Jautrumo sluoksnis atitinka tam tikro dalyko (nagrinėjamos temos – augmenijos, gyvūnijos, tradicinės kultūros) jautrumą, susijusį su modelio tikslu – išrinkti projektui tokią vietą, kurioje jo žala aplinkai ir kultūrai būtų kuo mažesnė. Į vieną jautrumo sluoksnį galima sujungti duomenis iš įvairių kraštovaizdžio objektus vaizduojančių GIS duomenų rinkinių – kelių, miško kirtaviečių arba gyvūnų buveinių.

Apribojimų žemėlapio kūrimo procesą sudaro šie pagrindiniai etapai:

- duomenų paruošimas arba modeliavimas (pvz., nuolydžio paviršiaus sudarymas arba gyvūnijos modeliai);
- apribojimų priskyrimas (įtakos buferių zonos, jautrumo reikšmės ir jungimo taisyklės);
- įvesties duomenų sujungimas ir jautrumo sluoksnių kūrimas;
- svorio daugiklių priskyrimas jautrumo sluoksniams;
- Jautrumo sluoksnių sujungimas į apribojimų žemėlapi.

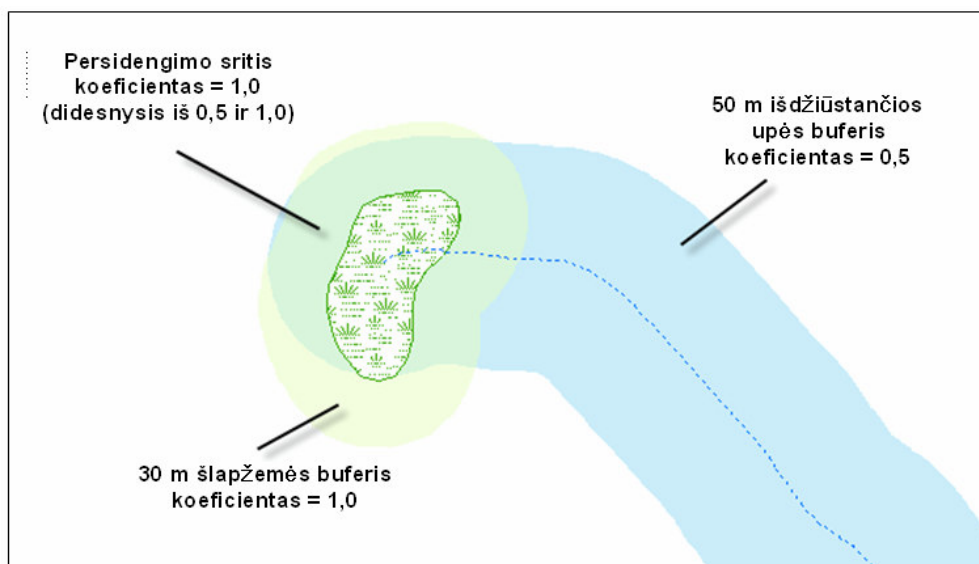
Duomenų paruošimas GIS

GIS duomenų rinkiniai apjungiami į grupes pagal jų apribojimų charakteristikas. Pavyzdžiui, šlapžemių duomenų rinkinio poligonai greičiausiai nurodo skirtingų tipų šlapžemių – aukštapelkių, žemapelkių, tarpinių pelkių ir liūnų – vietas. Jei šiems tipams priskiriami skirtingi svoriai, galima palikti atskirus tipus. Jei visiems tipams priskiriamas vienodas svoris, juos galima apjungti į vieną klasę (pvz., visos šlapžemės).

Kad GIS duomenų rinkinio elementą galima būtų naudoti modeliui, jam reikia priskirti jautrumo reikšmę – apribojimo koeficientą. Priskirta reikšmė atspindi konkretaus elemento jautrumą apribojimų atžvilgiu. Pavyzdžiui, jei vietos išrinkimo tikslas yra kiek įmanoma sumažinti projekto socialinę ir kultūrinę žalą, aktuali yra šlapžemių – ypač užliejamų pievų ir mišku apaugusių pelkių – svarba augmenijai ir gyvūnijai bei ekosistemos būklei. Todėl šioms šlapžemėms priskiriama gana didelė jautrumo reikšmė (labiau ribojanti projekto įgyvendinimą vietose, kuriose yra šių objektų). Apribojimų modelyje jautrumo reikšmės kito intervale nuo nulio (0,0) iki vieneto (1,0). Nulis reiškia, kad objektas nejautrus, o vienetas atitinka didžiausią jautrumą.

Jautrumo sluoksnių sudarymas

Jautrumo sluoksnis kuriamas iš vieno arba kelių GIS duomenų rinkinių. Kartais pasitaiko, kad skirtingų GIS duomenų rinkiniai su skirtingomis jautrumo reikšmėmis persidengia (5 pav.). Pavyzdžiui, kuriant „šlapžemių ir pakrančių“ jautrumo sluoksnį, 1,0 jautrumo šlapžemių buferis dažnai persidengia su 0,5 jautrumo išdžiūstančių upių buferiu. Jei tam pačiam jautrumo sluoksniui naudojami GIS elementai persidengia, į jautrumo sluoksnį įrašoma didžiausia jautrumo reikšmė. Šiuo atveju buvo įrašyta reikšmė 1,0.



5 pav. Koeficientų persidengimo pavyzdys

Ši „konservatyvi“ taisyklė galioja visiems GIS duomenų rinkiniams – į jautrumo sluoksnį įrašoma didžiausia iš visų jautrumo reikšmių.

Apribojimų žemėlapis sudarymas

Galutinis apribojimų žemėlapis gaunamas sujungiant jautrumo sluoksnius. Šie sluoksniai jungiami tiesinės svertinės sudėties būdu. Kiekvienam jautrumo sluoksniui priskiriamas svoris, nurodantis jo santykinę svarbą kitų jautrumo sluoksnių atžvilgiu. Šių svorių paskirtis – nurodyti jautrumo sluoksnio įtaką modeliui. Visų svorių suma turi būti 1,0 (arba 100 proc.). 5.1 lentelėje pateiktas pavyzdys, kaip galima priskirti svorius apribojimų žemėlapiui gauti naudojamiems jautrumo sluoksniams.

5.1 lentelė. Jautrumo sluoksnių svoriai

Jautrumo sluoksnis	Svoris
Šlapžemės ir pakrantės	0,492
Pagrindinės laukinių gyvūnų buveinės	0,268
Trikdžiai	0,140
Augmenija	0,053
Nuolydis	0,047
Suma	1,000

Kiekvieno apribojimų žemėlapių taško reikšmė skaičiuojama pagal formulę:

$$\text{apribojimų žemėlapių vieneto reikšmė} = (\text{šlapžemių ir pakrančių jautrumas} \times 0,492) + (\text{pagrindinių laukinių gyvūnų buveinių jautrumas} \times 0,268) + (\text{trikdžių jautrumas} \times 0,140) + (\text{augmenijos jautrumas} \times 0,053) + (\text{nuolydis} \times 0,047).$$

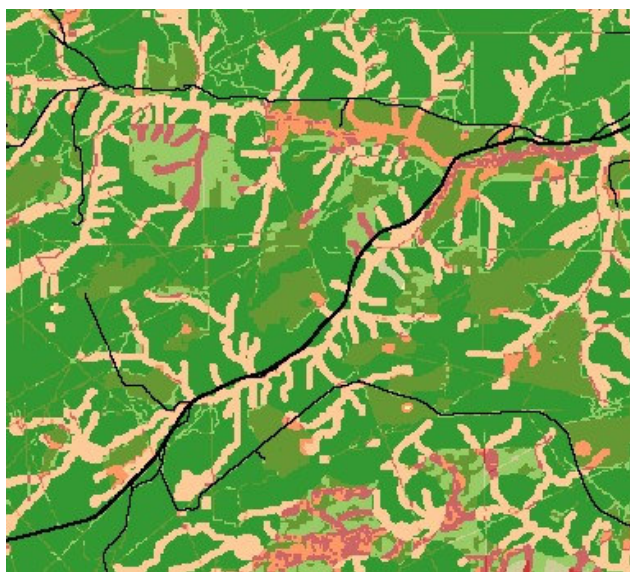
Toliau (5.2 lentelė) pateikiame žemėlapių vieneto (pvz. rastro gardelės) reikšmės skaičiavimo pavyzdį, kuriame gauta galutinė apribojimo reikšmė yra 0,707. Šiame

pavyzdyje jautrumo reikšmė yra sluoksnio jautrumo reikšmė konkrečioje sluoksnio gardelėje. Svertinis jautrumas skaičiuojamas kiekvienam tiriamos srities žemėlapiu vienetui. Svertinio jautrumo reikšmės yra intervale nuo 0,0 (apribojimų nėra) iki 1,0 (visiškai draudžiama). Apribojimo reikšmė yra svertinių jautrumo reikšmių suma konkrečiame žemėlapiu vienetu.

5.2 lentelė. Apribojimų skaičiavimo pavyzdys

Jautrumo sluoksnis	Jautrumo reikšmė	Svoris	Svertinis jautrumas
Šlapžemės ir pakrantės	1	0,492	0,492
Pagrindinės laukinių gyvūnų buveinės	0	0,268	0,000
Trikdžiai	1	0,140	0,140
Augmenija	0,8	0,053	0,042
Reljefas	0,7	0,047	0,033
		SUMA (apribojimo reikšmė)	0,707

5.6 pav. pateiktas apribojimų žemėlapiu pavyzdys. Rudos sritys ribojamos labiausiai, o tamsiai žalios – mažiausiai. Šiuo atveju akivaizdu, kad šlapžemių ir pakrančių įtaka galutiniam apribojimų žemėlapiu yra didžiausia – gautame paviršiuje aiškiai matyti vandens objektų buferiai.



5.6 pav. Apribojimų žemėlapiu pavyzdys

Prielaidos ir trūkumai

Dėl modelio trūkumų apribojimų žemėlapius reikia interpretuoti atsargiai. Didžiausias trūkumas yra tas, kad galutiniam žemėlapyje nebelieka atskirų (priskirtų atskiriems kraštovaizdžio objektams) jautrumo reikšmių. Kadangi rezultatas apjungia informaciją iš įvairių GIS duomenų šaltinių, atskirų elementų jautrumas sumažinamas. Tai atsitinka dėl dviejų priežasčių.

- Iš visų įvedamų GIS duomenų jautrumo reikšmių konkrečioje kraštovaizdžio vietoje į jautrumo sluoksnį įrašoma didžiausia reikšmė. Tai reiškia, kad jei vienoje srityje bus trys objektai, ir visų jų jautrumas bus 1,0, ši sritis atrodys ne svarbesnė negu sritis su vienu 1,0 jautrumo objektu. Abiems sritims bus priskirtas 1,0 jautrumas.
- Galutinio apribojimo reikšmės tikslumą mažina ir jautrumo sluoksniams sujungti naudojamas svorių mechanizmas. Jei, pavyzdžiui, šlapžemių ir pakrančių jautrumo svoris būtų daug mažesnis (pvz., 0,10), žymiai padidėtų kitų sluoksnių įtaka galutiniam rezultatui, ir apribojimų žemėlapis gerokai pasikeistų. Santykinė šlapžemių ir pakrančių svarba (ar nesvarba) galutiniame žemėlapyje neatsispindėtų.

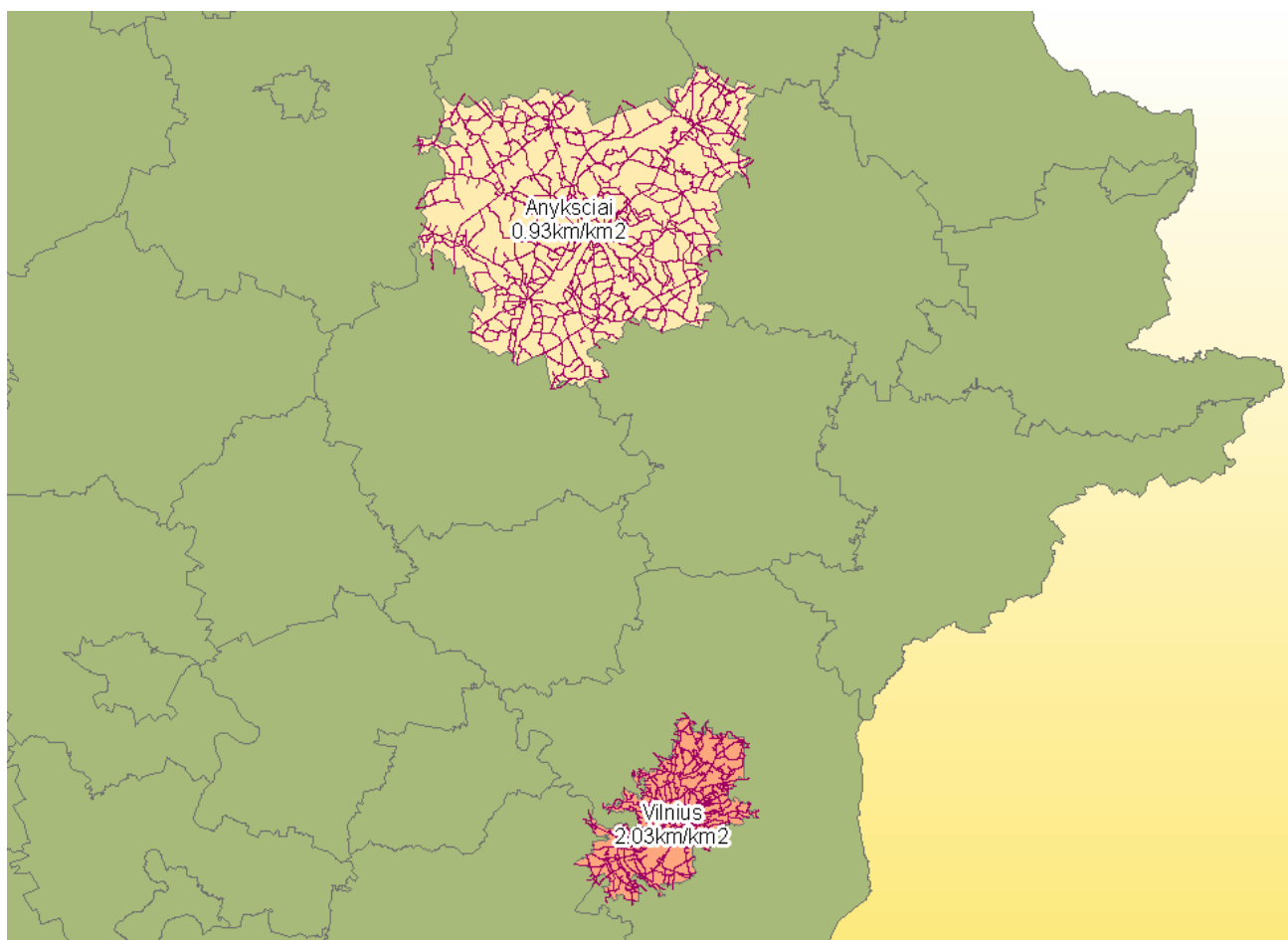
Taikymas

Kaip matyti iš išnagrinėto pavyzdžio, apribojimų žemėlapiai gali suteikti informacijos apie aplinkosauginius apribojimus, bet jie tinka ir galimoms projekto vietoms nustatyti. Idealiu atveju būtų atliekama dviejų rūšių analizė: apribojimų analizė nustatytų jautrius aplinkos elementus, o galimybių analizė tirtų esamas plėtros ir vietovės sąlygas ir nustatytų su konkrečia vieta susijusius „teigiamus“ veiksnius (pvz., kelio per artimas žemes teises, mažą nuolydį, mažiausią kelio sankirtų su upėmis skaičių). Šių dviejų skaičiavimų rezultatus galima būtų naudoti ir atskirai, ir kartu – viename žemėlapyje, vaizduojančiame optimalias ir inžinerinių reikalavimų, ir aplinkosaugos požiūriu projekto vietas. Pagrindinis apribojimų tikslas – pateikti informaciją, padėsiančią priimti geriau pagrįstus sprendimus ir taip sumažinti neigiamą pramonės plėtros poveikį. Apribojimų žemėlapiai leidžia efektyviai nustatyti vietas, kuriose veiklos daroma žala būtų maža; ir nors šias vietas vis tiek reikia patikrinti lauko tyrimais, apribojimų žemėlapiai leidžia sutelkti lauko tyrimų programas į mažesnes vietovės dalis.

5.5.3 Kelių tankis

Kelių tankis yra naudingas aplinkos rodiklis, padedantis įvertinti dabartinę ir būsimą žmogaus veiklos sukeltą trikdžių įtaką laukinių gyvūnų ir žuvų buveinėms. Juo remiantis galima išmatuoti veiklą tam tikroje srityje ir buveinių fragmentacijos laipsnį. Paprastai kelių tankis išreiškiamas kaip kelių ilgis kilometrais viename kvadratinėje kilometre, arba, jei reikia lyginti skirtingas sritis – visoje srityje, pavyzdžiui, vandens baseine arba administraciniame vienetu (5.7 pav.). Jei skaičiuojamas skirtingų laiko periodų tankis (pvz., praeito amžiaus devintojo, dešimtojo ir šio amžiaus pirmojo dešimtmečio), gautus rezultatus galima palyginti ir nustatyti laikines vystymosi tendencijas ir galbūt nustatyti perkrautas sritis.

Atliekant kelių tankio analizę, galima laikyti, kad visų kelių įtaka tankio statistikai yra vienoda, arba platesniems ar intensyvesnio eismo keliams priskirti didesnius svorius ir apskaičiuoti svartinį kelių tankį.



5.7 pav. . Kelių tankis Anykščių rajone ir Vilniuje

5.5.4 Bendroji dirvožemio nuostolių formulė

Bendroji dirvožemio nuostolių formulė (*Universal Soil Loss Equation* – USLE) yra įprastinis dirvos erozijos modelis, įvertinantis klimato charakteristikas, dirvožemio savybes, topografiją, paviršiaus sąlygas ir žmonių veiklą. Jis numato vidutinius dirvožemio nuostolius dėl nuoslinkio nuolaidžiu paviršiumi žemdirbystės, ganyklų bei kitose prižiūrimose žemės sistemose (pvz., statybvietėse).

Formulė atrodo taip: $A = R K L S C P$,
kur

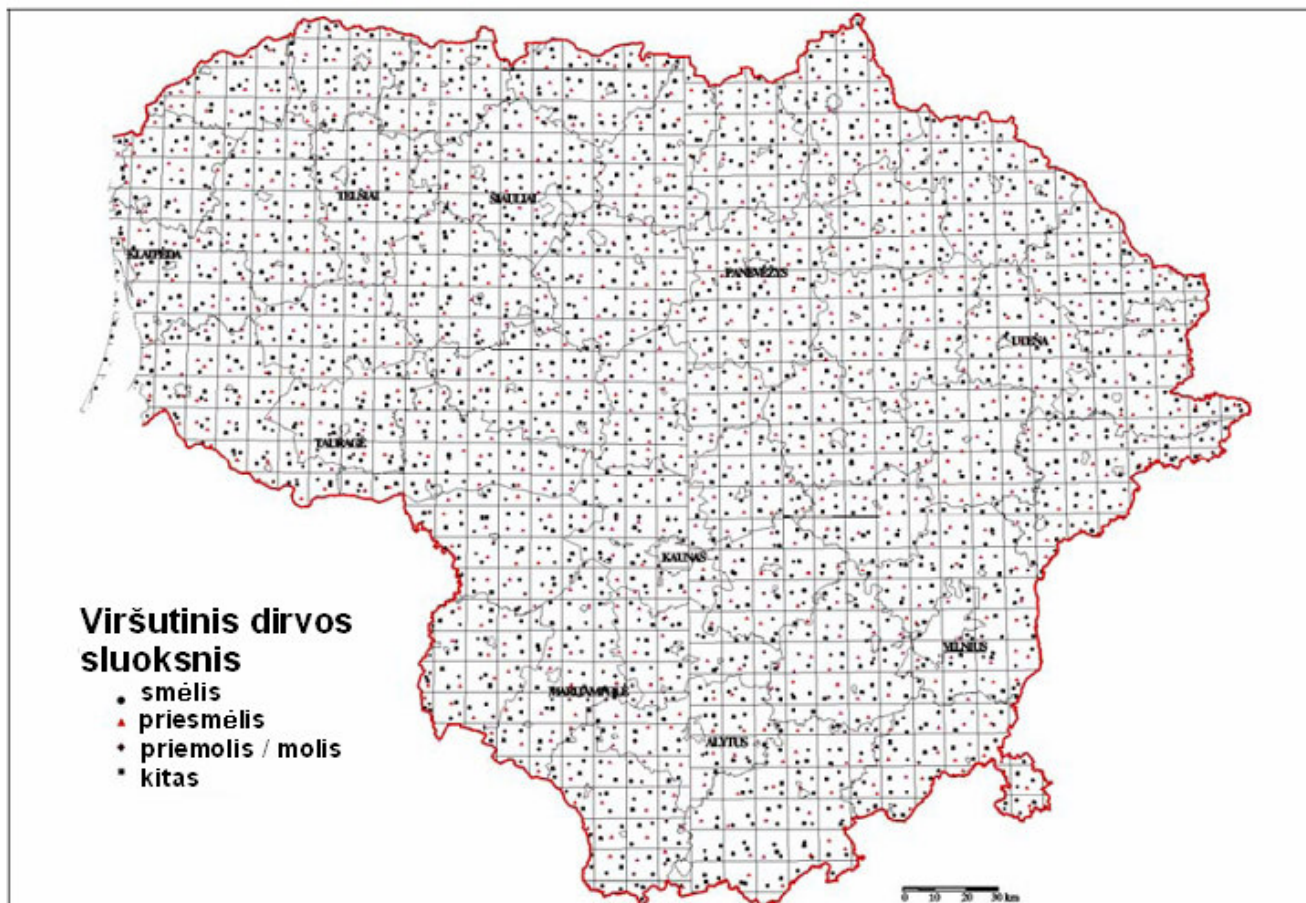
- A** – vidutiniai dirvožemio nuostoliai (pvz., tonomis vienam akrui per metus);
- R** – lietaus ardomas poveikis (apskaičiuojamas iš vidutinės lietaus energijos);
- K** – dirvožemio erozijos koeficientas (vidutiniai dirvožemio nuostoliai tonomis vienam akrui per metus, esant standartiniam šlaito ilgiui ir statumui);
- L** – šlaito ilgis;
- S** – šlaito statumas;
- C** – agrotechninis koeficientas (agrotechninių veiksnių įtaka dirvos erozijai);

P – tvirtinimo koeficientas (įvertina kontūrinę ar juostinę žemdirbystę, šlaitų terasas, požeminį drenažą).

Visi šie veiksniai įvertinami modeliuojant dirvos erozijai įtakos turinčias sąlygas konkrečioje vietoje.

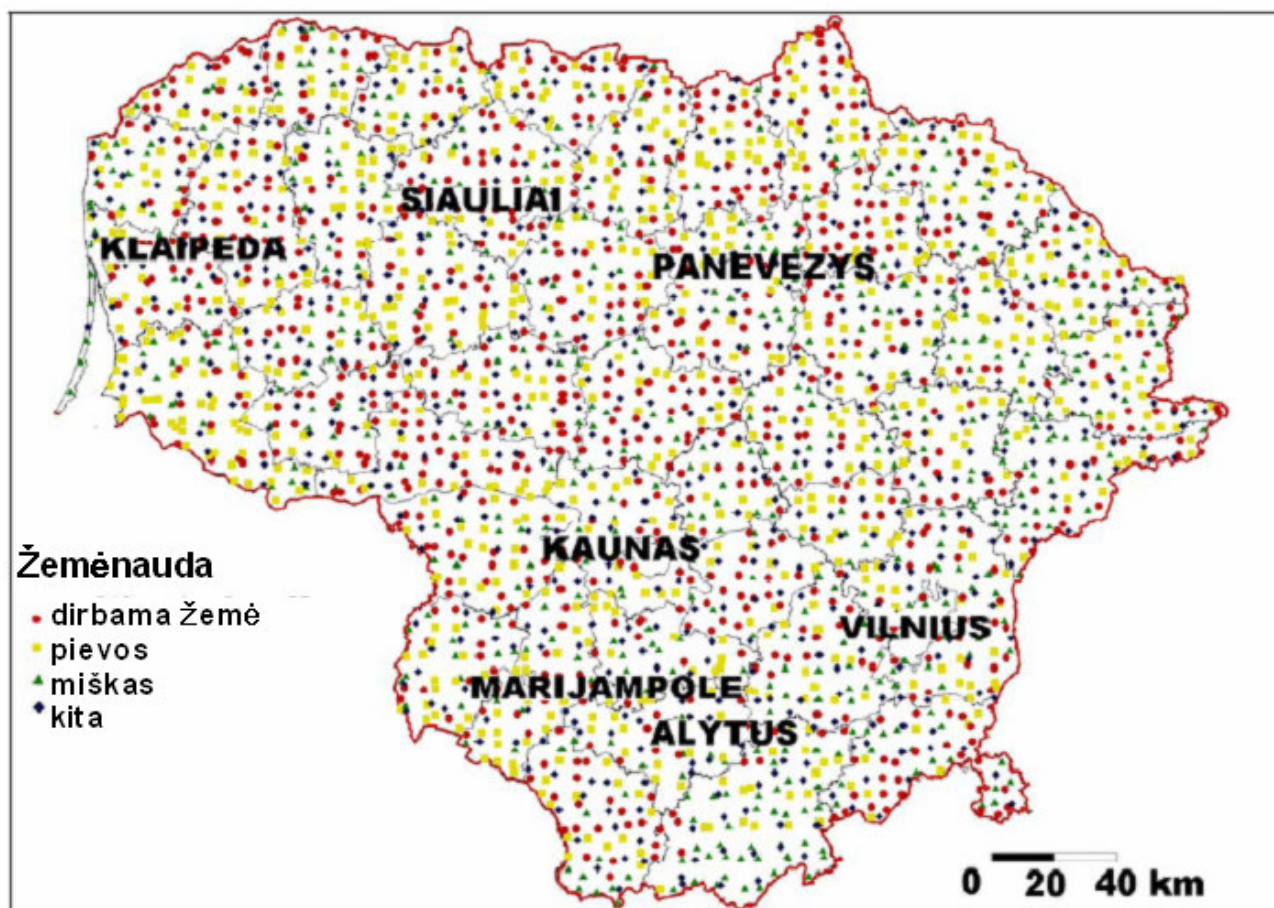
GIS leidžia modeliuojant atsižvelgti į erdvinę lygties dalį – nustatyti:

- dirvožemio nuoslinkį dėl lietaus ardomojo poveikio (R) – pagal kritulių duomenis,
- dirvožemio erozijos koeficientą (K) – pagal dirvožemio savybių žemėlapi (5.8 pav.),



5.8 pav. Lietuvos dirvožemio savybių žemėlapis

- gardelių šlaito ilgį (L) ir statumą (S) – iš skaitmeninių aukščių modelių,
- agrotechninį koeficientą (C) ir tvirtinimo koeficientą (P) galima nustatyti iš žemės dangos ir (arba) žemėnaudos duomenų (5.9 pav.).



5.9 pav. Lietuvos žemėnaudos žemėlapis

Iškelti ir paruošti įvesties duomenys perdangos operacija perklasifikuojami į ekvivalentiškus vienetus, sujungiami pagal bendrąją dirvožemio nuostolių formulę, ir gaunamas dirvos erozijos potencialo žemėlapis.

Dalies klausimai savarankiškam darbui:

- Kas yra modelis, ir kokie yra modelių trūkumai?
- Kaip galima patikrinti modelio rezultatus?
- Kas yra dvejetainis modelis? Kokio pobūdžio uždaviniams spręsti jis tinka?
- Kas yra 0 skyriuje aprašytas kelių tankio modelis – aprašomasis, aiškinamasis, prognozuojamasis ar norminamasis? Kodėl?
- Ar 0 skyriuje pateikta bendroji dirvožemio nuostolių formulė yra vaizduojamasis, ar proceso modelis? Kodėl?

Literatūra

1. Blyth, C. Ann, and David Cake, *Constraints Mapping as a Decision Support Tool for Project Siting*. GeoTec 2005 Conference. Vancouver, BC, February 13-15, 2005.
2. Chang, K.T. *Introduction to Geographic Information Systems*. McGraw-Hill, 2006.
3. Chrisman, N. *Exploring Geographic Information Systems*. John Wiley and Sons, 1997.
4. Heywood, I., Cornelius, S. and Carver, S. *An Introduction to Geographical Information Systems*. Pearson Education, 2002,