

KLAIPĖDOS UNIVERSITETAS  
Gamtos ir matematikos mokslų fakultetas  
Statistikos katedra

BALTIJOS JŪROS SKAITMENINIO DUGNO RELJEFO SUDARYMAS STOCHASTINIAIS IR  
DETERMINISTINIAIS METODAIS  
Baigiamasis magistro darbas

Autorius SOT-1 stud. Tamara Strelkauskytė  
Vadovas dr. Ingrida Borisenko

**KLAIPĖDA, 2012**

## TURINYS

<b>VADAS</b>	7
<b>I. DETERMINISTINIAI IR STOCHASTINIAI METODAI</b>	9
1.1 <i>Skirtumai tarp stochastini ir deterministini modeli</i>	9
1.2 <i>Stochastiniai modeliavimo metodai – krigingas</i>	13
1.3 <i>Deterministiniai metodai</i>	17
1.4 <i>Krigingas ir splain charakteristik palyginimas</i>	23
<b>II. STOCHASTINI BEI DETERMINISTINI METOD PANAUDOJIMO MATEMATIN INTERPRETACIJA</b>	25
2.1 <i>Splain funkcijos bei j panaudojimas</i>	25
2.2 <i>Splain erdv</i>	28
2.3 <i>Pirmos eil s splaino s vyb s</i>	30
2.4 <i>Ermito splainas</i>	30
2.5 <i>B – splainas</i>	32
2.7 <i>Ordinarusis krigingas</i>	36
2.8 <i>Paprastasis krigingas</i>	37
2.9 <i>Universalus krigingas</i>	39
2.10 <i>Medianinio išlyginimo krigingas</i>	40
2.11 <i>Modeli patikros metodai bei analiz</i>	42
<b>III. STOCHASTINI BEI DETERMINISTINI METOD PALYGINIMAS NAUDOJANT R PROGRAMA</b>	44
3.1 <i>Baltijos j ros gyli skaitmeninio reljefo sudarymas stochastiniais metodais</i>	44
3.1.1 <i>Universalus krigingo pritaikymas</i>	44
3.1.2 <i>Medianinio išlyginimo krigingo pritaikymas</i>	49
3.2 <i>Baltijos j ros gyli skaitmeninio reljefo sudarymas deterministiniais metodais</i>	53
<b>LITERAT ROS S RAŠAS</b>	59
<b>PRIEDAI</b>	61

## SANTRAUKA

*Tamara Strelkauskytė. Baltijos jūros skaitmeninio dugno reljefo sudarymas stochastiniais ir deterministiniais metodais: statistikos magistro darbas / Mokslinio darbo vadovė dr. Ingrida Borisenko. Klaipėdos universitetas. Gamtos ir matematikos mokslų fakultetas. Statistikos katedra. - Klaipėda, 2012. – 87 p.*

*Raktiniai žodžiai:* stochastiniai metodai, deterministiniai metodai, krigingas, splainas, Baltijos jūros gyliai.

Magistro baigiamojo darbo tikslas patvirtinti arba paneigti iškeltą hipotezę, kad taisyklingai išsidėsiusiems Baltijos jūros gylių duomenims abiejų taikytų metodų (krigingo ir splaino) tikslumas bus panašus.

Darbe iškeltai hipotezei patikrinti analizuojami trisdešimt Kuršių marių bei Baltijos jūros gylių duomenys, kurie gauti iš Klaipėdoje esančio „Jūrinio tyrimo centro“. Šiems duomenims kringingo bei splaino metodais atliekama interpoliacija bei randamas skaitmeninis dugno reljefo paviršius. Baltijos jūros ir Kuršių marių skaitmeniniam paviršiui rasti naudojami R programos `gstat`, `spatial`, `Akima`, `mba`, `fields`, `gss` paketai.

Skaitmeninis Baltijos jūros dugno paviršius yra vaizduojamas atliekant interpoliaciją universaliojo, medianinio išlygimo kringingu, tiesiniu, kubiniu, kubiniu Akima bei kubiniu B – splainu. Interpoliacijos tikslumui garantuoti yra tiriami: trendo tipas, semivariogramos laipsnis, tipas, priklausomybė tarp gylių ir koordinatės. Tikslumui vertinti naudojamas Akaike informacijos kriterijus bei mažiausi kvadratų metodas. Atlikus interpoliaciją geriausias prognozavimo metodas randamas: atlikus kryžminį patikrą (palyginus vidutinį kvadratinį paklaidos, šaknies iš vidutinio kvadratinio prognozės paklaidos bei vidutinio kvadratinio nuokrypio reikšmes); bei skaitiniais metodais atlikus anova analizę (palyginus F kriterijaus  $p$  reikšmes).

Atlikus interpoliaciją bei skaitiniais metodais palyginus gautus rezultatus padaryta išvada, kad Baltijos jūros rytinės dalies ir Kuršių marių vakarinės dalies teritorijai interpoliuoti yra tinkamiausi: universalusis kringingas (teigiant, kad duomenyse egzistuoja trejų laipsnio trendas) bei kubinis splainas. Gautas modelių rezultatų tikslumas bei reljefo paviršius abejais atvejais praktiškai identiškas, todėl patvirtinama baigiamajame darbe iškelta hipotezė, kad taisyklingai išsidėsiusiems duomenims abiejų taikytų metodų tikslumas panašus.

## SUMMARY

*Tamara Strelkauskyt . Digital terrain modeling of Baltic sea bottom with stochastic and deterministic methods: statistical master's thesis / academic work, Dr. Ingrida Borisenko. Klaip da University. Sciences and Mathematics Faculty. Department of Statistics. - Klaip da 2012. – 87 p.*

*Keywords:* stochastic methods, deterministic methods, kriging, spline, the Baltic sea depths data.

The aim of this master thesis a stochastic forecast data and deterministic interpolation methods and compare them. Also, to confirm or refute the hypothesis for the Baltic Seas depth arrangement data accuracy between the two methods used is similar.

The study analyzed 30 of the Curonian Lagoon and the Baltic Sea depth data obtained from the „Sea Research Center“ in Klaip da . From this data, interpolation was performed by numerical methods and found the bottom of a digital terrain surface. To find the digital surface used gstat R program, Spatial, Akima, mba, fields, gss packages.

Digital Baltic seabed surface is visualised and represented by using the interpolation: universal kriging, median polishing kriging and linear, cubic, cubic Akima and cubic B – spline methods. Interpolation is investigated to ensure the accuracy of the data: the trend type, semivariogram degree, type, and the relationship between the depth and coordinates. Used to assess the accuracy of akaike's information criterion and sum of square of errors methods. The best interpolation method is found using: the cross-validation (comparing mean square error, root mean square of forecast errors and standard deviation values), and numerical methods using ANOVA analysis (F-test comparing p values).

After interpolation, and the numerical methods used for comparing the obtained results it was concluded that the Baltic Sea in the eastern part and western part of the Curonian Lagoon area is interpolated best with universal kriging (stating that the data exists in the third degree trend) and a natural cubic spline. The resulting models of the accuracy and surface relief in both cases was almost identical, and therefore confirms the hypothesis that the correct data arrangements, both the accuracy of the methods is similar.

## LENTELI S RAŠAS

1.1 lentel . Deterministini ir stochastini metod suskirstymas	10
2.1 lentel . Kubinio B - splaino reikšm s pateiktuose taškuose	33
3.1 lentel . Analizuojamos Baltijos j ros teritorijos koordinat s	44
3.2 lentel . ANOVA analiz s rezultatai, Akaik s informacijos kriterijaus (AIC) reikšm	46
3.3 lentel . Tre iojo laipsnio trendo polinomin s lygties koeficientai	46
3.4 lentel . Semivariogram modeliavimo funkcijos	47
3.5 lentel . Priglodint semivariogram kvadratiniai nuokrypi sumos.	48
3.6 lentel . Duomenys apie tinkl	49
3.7 lentel . Pagrindines liekan statistin s charakteristikos.	50
3.8 lentel . Akaik s informacijos kriterijaus reikšm .	50
3.9 lentel . Semivariogram modeliavimo funkcijos	51
3.10 lentel . Priglodint semivariogram kvadratiniai nuokrypi sumos	52
3.11 lentel . Kryžmin s patikros rezultatai	53
3.12 lentel . Koordina i kolinearumo diagnostika	54
3.13 lentel . Modelio tinkamumo reikšm	55

## PAVEIKSLŲ RAŠAS

1.1 pav. Aštrūs kampai tiesinio splaino viršūnėse	18
1.2 pav. a) Kubinio splaino kreivė, b) akima splaino kreivė	19
1.3 pav. a) Akima ir b) kubinio splaino grafinis palyginimas	20
1.4 pav. a) B – splain kreivė, b) Bezier kreivė	20
1.5 pav. B – splainai turintys skirtingą eilę	21
1.6 pav. IDW metodo panaudojimo kriterijus	22
2.1 pav. Dviejų bazinių, dvipusių funkcijų grafinis vaizdas	35
3.1 pav. Baltijos jūroje išmatuoti gyliai bei jų koordinatų ekstremumai	44
3.2 pav. a) Logaritmuoti, b) nelogaritmuoti gylių duomenų histogramos	45
3.3 pav. a) Semivariograma, pašalinus I – ojo laipsnio trendą b) jos žemėlapis	45
3.4 pav. Kryptingos semivariogramos	46
3.5 pav. Priglodintas a) sferinis, b) eksponentinis, c) Gauso modelis	47
3.6 pav. a) Universaliojo krigingo prognozių rezultatai b) prognozių paklaidos	48
3.7 pav. a) Nelogaritmuoti, b) logaritmuoti gylių duomenų histogramos atitinkamai	51
3.8 pav. Liekanų semivariograma, kai vidurkis pastovus bet nežinomas	51
3.9 pav. Kryptingos semivariogramos liekanų duomenims.	51
3.10 pav. Priglodintas a) eksponentinis, b) sferinis, c) Gauso modelis	52
3.11 pav. a) Liekanų prognozės bei b) galutinų gylių duomenų prognozės	52
3.12 pav. a) Tiesinė ir b) kubinė interpoliacija splainais pritaikyta Baltijos jūros gyliams	54
3.13 pav. a) Kubinis Akima bei b) kubinis B - splainas pritaikytas Baltijos jūros gyliams prognozuoti	55

## VADAS

Geostatistinis paviršiaus modeliavimas – yra nauja paviršiaus modeliavimo šaka. Joje taikomi sudingti paviršiaus modeliavimo algoritmai bei statistiniai matavimo duomenų apdorojimo metodai. Uždaviniai su kuriais susiduria geostatistika išsiskiria tuo, kad pradiniai matavimo rezultatai yra tarpiniai tarp grynai atsitiktini ir grynai determinuot dydžių. Sprendžiant tokius uždavinius reikia pasinaudoti regionalizuotais duomenimis, kurių stochastinės savybės yra teritorijos geografinės padėties funkcijos. Šios srities tyrimai prasidėjo jau XIX amžiaus Alexander von Humboldt darbais ir tęsiasi iki šiol. Šiomis dienomis toks modeliavimas apima mokslo šakas apie žemėlapių geoinformatiką ir geostatistiką – ir turi daug pavadinimų. Pavyzdžiui vakaruose toks modeliavimas vadinamas kiekybinis reljefo analizė, geomorfometrija arba kiekybinis geomorfologija. Duomenys skaitmeniniam reljefo modelio sudarymui yra gaunami naudojantis keliomis priemonėmis. Duomenys gali būti renkami skaitmeniniu pavidalu, siekiant tiesiogiai panaudoti geografinės informacinės sistemas (GPS matavimai, skaitmeninis aerofotogrametrija ir kosminis nuotrauka, skaitmeninis topografinis nuotrauka) arba tai gali būti skaitmeniniai ir analoginio pavidalo duomenų rinkiniai, surinkti kitais tinklais ir konvertuoti reikiamam pavidalui. Duomenų rinkimas yra vienas brangiausių ir didžiausių laiko išlaidų reikalaujantis etapas. Kai kurie šaltiniai teigia, kad duomenų rinkimas gali sudaryti iki 85 procentų geografinės informacinės sistemos kaštą. Dideli finansiniai sunkumai ir dideli kitoki priežastiniai gali atsirasti vietose, kuriuose matavimai nebuvo atlikti. Sudarant skaitmeninius reljefo modelius distanciniais matavimo metodais iš kosmoso arba dirbtinio žemės palydovo, susiduriama su paviršiaus fragmentais, kuriuose dideli tankios augalijos arba kitoki priežastiniai paviršiaus išmatuoti ne manoma, o juose atliekami matavimai tik tam tikrose pastatytose stotyse (jų kiekis ribojamas dideli finansiniai priežastiniai). Taigi realiose situacijose dažnai atsitinka, kad surinkta informacija nėra tolydžiai pasiskirsčiusi analizuojamos teritorijos ribose arba turimų duomenų kiekis nėra pakankamas skaitmeninio reljefo paviršiaus sudarymui. Siekiant atvaizduoti paviršiaus reljefą, svarbu turėti tankių vietovės taškų tinklą su koordinatėmis ir gyliais (ar kitokia charakteristika). Šiam tikslui, remiantis atramos taškų duomenimis, taikomi vairūs interpoliavimo metodai. Kuo daugiau šių taškų, tuo interpoliavimo rezultatai patikimesni. Pagrindinės charakteristikos, turimos tokos, skaičiuojant nežinomus paviršius yra modeliuojamas plotas, paviršiaus forma, erdvinis išsidėstymas. Keičiantis bet kuriai iš šių charakteristikų, keisis ir nežinomosios reikšmės. Nustatius prognozavimui reikalingas charakteristikas yra parenkamas vienas iš interpoliavimo metodų – stochastinis arba deterministinis (krigingas arba splinas atitinkamai).

Šiame darbe bus naudojami Baltijos jūros ir Kuršių marių duomenys. Teisingai pasirinktas interpoliavimo metodas, nagrinėjamiems duomenims, padeda surasti ryšius tarp jau žinom

reikšmi ir prognozuoti tarp jų esančias bei taip sudaryti pakankamai tikslų, todėl jų rodo dugną. Informacija apie reljefą labai svarbi ir naudojama sprendžiant vairius projektavimo, telekomunikacijų, hidrotechninius uždavinius, karo inžinerijoje ir kitur (gyvenamųjų kvartalų statymas nepradedamas neišanalizavimus reljefo paviršiaus, augalijos pasiskirstymo d sningumus nulemia dr gnis bei šiluma, žuvingumas priklauso nuo gylio ir t.t).

Kaip yra žinoma papildom matavimo stoties rengimas, reikiam matavimo agregat pirkimas ar nuoma reikalauja didelių išlaidų. Baltijos jūros dugno interpoliavimas atliekamas siekiant sumažinti resursus bei išlaidas naudojamas atliekant nežinom charakteristik , šiuo atveju gylio matavimus. Šiame darbe tolydus Baltijos jūros dugno paviršius sudaromas naudojantis stochastiniais bei deterministiniais interpoliavimo metodais.

**Darbo tikslas** atlikus interpoliaciją patvirtinti arba paneigti hipotezę, kad taisyklingai išsidėsiusiems gylio duomenims tinkamai pasirinkt skirting taikyt metod tikslumas bus panašus.

Tam, kad būtų gyvendintas darbo tikslas yra būtina atlikti šiuos uždavinius:

- Aprašyti ir palyginti stochastinius ir deterministinius interpoliavimo metodus bei išsiaiškinti su metodais susijusias savybes.
- Pateikti bei išanalizuoti pagrindini bei dažniausiai naudojam metod matematines interpretacijas.
- Pasinaudojus R programos atitinkamais paketais atlikti Baltijos jūros gylio prognoz (pasinaudojus duomenims tinkamais keliais stochastiniais ir keliais deterministiniais metodais).
- Naudojantis modeli patikros metodais palyginti stochastiniais bei deterministiniais interpoliavimo metodais gautas prognozes (t.y skaitmeninius reljefo paviršius).

Šiame darbe pirmajame skyriuje yra aprašyti bei palyginti deterministiniai bei stochastiniai metodai (krigingas, splainas, atvirkštinio atstumo metodas). Antrajame pateikta jų matematin interpretacija. Trečiajame skyriuje, pasinaudojus R programos paketais, atlikta duomen interpoliacija, randamas skaitmeninis reljefo paviršius (keliais iš aprašyt metod ), nustatomas rezultat tikslumas, bei atliekamas metod palyginimas.

Darbe pateikta 14 lentelė bei 20 paveikslai.

# I. DETERMINISTINIAI IR STOCHASTINIAI METODAI

## 1.1 Skirtumai tarp stochastini ir deterministini modeli

Modeliavimas yra neatsiejama bet kurio mokslo dalis - tiek socialinio, tiek gamtos. Realios pasaulio sistemos paprastai yra labai sudėtingos. Norint šias sistemas suprasti, prognozuoti jų elgesį ar kontroliuoti, būtina jas supaprastinti, o tai reiškia sukurti modelį. Kuriant modelius yra naudojami keli metodai, pavyzdžiui reljefo paviršius gali būti modeliuojamas stochastiniais ir (arba) deterministiniais metodais. Turint modelį, pagal turimus taškus, sudaromas tolydus paviršius. Technika, kuri tai atlieka vadinama interpoliacija, tačiau, kalbant griežtai, tai galioja tik taškams, geografiškai patenkantiems ribas, apibrėžiamas atrankos taškais (priešingu atveju tai yra ekstrapoliacija). Kalbant apie geostatistiką, šios technikos gali būti laikomos interpoliatoriumi, glodintuvu, vertinimo arba prognozės rankiu, atsižvelgiant konkrečią matematinę ir (arba) statistinę techniką. Taigi sudarius modelį atskiri taškai yra apjungiami taip, kad gauti tinkamus duomenis vienas tarp duotų taškų, o vertinami kintamieji yra esantys jėm reikšmių svertiniai vidurkiai. Jie yra naudojami stochastinės interpoliacijos metu.

Stochastinė interpoliacija prognozėms atlikti naudoja tikimybinus modelius. Tai tokie modeliai, kuriame pradiniai parametrai ir modelio parametrai gali gyti atsitiktines reikšmes ir modeliuojama atsitiktinomis lygomis. Atsitiktinio lauko modelis šiuo atveju

$$Z(\mathbf{s}) = \tilde{z}(\mathbf{s}) + v(\mathbf{s}) \quad (1.1)$$

čia  $Z(\mathbf{s})$  – interpoliuojamas kintamasis (pavyzdžiui, azoto oksido koncentracija, reljefas);  
 $\tilde{z}(\mathbf{s})$  – deterministiniais metodais nustatomas trendas;  
 $v(\mathbf{s})$  – autokoreliuota erdvinė atsitiktinė paklaida. „Autokoreliuota erdvinė“ reiškia, kad nors svyravimai iš anksto nėra žinomi, tačiau artimose padėtyse jie yra panašesni savo reikšmėmis nei tolimesnėse.

$\{v(\mathbf{s}) : \mathbf{s} \in D\}$  - klaidos atsitiktinis laukas matavimo vietose  $\mathbf{s}$ .

Taip pat stochastiniai interpoliacijos metodai šiuo prognozės paklaidos  $v(\mathbf{s})$  vertinimu su numatytais skirtumais. Ši grupė apima kriginę ir visus antrinius metodus tokius kaip trendo paviršiaus vertinimas, regresija ar kokrigingas (Blyth et al., 2008). Stochastinis modelis, kuris taip pat dar vadinamas tikimybinu, vaizduojamas matematine lygtimi, kurios bent vienas parametras tam tikru laipsniu yra atsitiktinis. Dėl šio atsitiktinumo atsiranda tam tikra modelio rezultato paklaida  $v(\mathbf{s})$  ( $Z(\mathbf{s}) = \tilde{z}(\mathbf{s}) + v(\mathbf{s})$ ) arba neapibrėžtumas. Neapibrėžtumas yra būdingas visiems pavojų ir rizikos vertinimams, kuriuose matematiniai modeliai naudojami apibūdinti netiesiogiai stebimų informacijų. Šis uždavinys kyla dėl to, kad sudaromi modeliai nepilnai atspindi realybę, kad ne manoma visiškai tiksliai nustatyti pradinio modelio parametrų reikšmių ir realiose situacijose jos

gali būti skirtingos, nei naudojamos modelyje. Jis yra laikomas neišvengiamu ir išreiškiamas tikimybe. Kaip pavyzdį galima pateikti modelį, kuris vertina nuošliaužų tikimybę, jei būtų iškirstos tam tikros miško sritys. Dar vienas pavyzdys yra kriginio analizės taikymas vandens m gini matavimams. Atlikus šį procedūrą gautume interpoliuotus paviršius, kurie vaizduoja galimą tam tikro teršalo koncentraciją bei standartines visų numatomų reikšmių paklaidas. Pagrindinis skirtumas tarp stochastinių ir deterministinių modelių yra tai, kad deterministiniame modelyje nei vienas iš matematinės lygties kintamųjų arba parametrų nėra laikomas atsitiktiniu.

Deterministiniai metodai bando pritaikyti taškams, turintiems tam tikras  $x$ ,  $y$  ir  $z$  koordinates, vien iš matematinio paviršiaus tipą. Deterministiniais interpoliavimo metodais duomenų gavyba – tai duomenų apdorojimas, naudojant sudėtingas duomenų paieškos galimybes ir statistinius algoritmus pasikartojančių šablonų ir koreliacijų radimui turimose didelėse duomenų bazėse; naujų duomenų prasmės aptikimo, nustatymo, atradimo būdas. Centrinei tendencijai arba viduriui nusakyti yra naudojami tokie matai kaip vidurkis (vidutinio požymio reikšmė, nustatyta tiriant skirtingus objektus), mediana (požymio reikšmė, kuri dalija variacinę eilutę, populiaciją ar tikimybių skirstinį dvi lygias dalis) ir moda (tai dažniausiai pasikartojanti požymio reikšmė imtyje). Tokie modeliai kuriami tam, kad juos būtų galima naudoti erdviniai sprendimai priimimo sistemose, jie leidžia vartotojui eksperimentuoti su pasaulio kopija, bet ne su realiais objektais bei leidžia operatyviai pateikti varias galimas realias situacijas alternatyvas. Šiai grupei priskiriami: atvirkščiai proporcingo atstumo metodas, globalūs ir lokalūs polinomialai bei spindulio tipo funkcijos ar splineai. Jie dažniausiai naudojami glotniems formų paviršiams interpoliuoti, taip pat tolygiai kintančio paviršiaus nuolydžiui apskaičiuoti.

### 1.1 lentelė. Deterministiniai ir stochastiniai metodų suskirstymas

Globalūs		Lokalūs	
Deterministiniai	Stochastiniai	Deterministiniai	Stochastiniai
<b>Trendo paviršius</b> (netikslus)	<b>Regresija</b> (netikslus)	<b>Atvirkščiai proporcingo atstumo metodas</b> (tikslus) <b>Splineai</b> (tikslus)	<b>Krigingas</b> (tikslus)

Kaip matome iš lentelės (1.1) deterministinius bei stochastinius metodus dar galima suskirstyti globalius bei lokalius. Lokalūs modeliai, kurie yra tikslūs, naudoja tik informaciją iš tam tikro ploto. Šis plotas gali būti keičiamas kol interpoliavimo rezultatais padengiamas visas paviršius. Priešingas lokaliui yra globalus interpoliavimo metodas, kuris interpoliavimo proceso metu atsižvelgia visus duomenų taškus, nepriklausomai nuo jų padėties, todėl šis metodas yra laikomas netikslu.

Jeigu metodas vadinamas globaliu, tuomet iš viso duomen rinkinio gaunamas vienas statistinis rezultatas rinkinys ir jis aprašo visą tiriamą sritį. Globaliems metodams yra priskiriami trendas bei regresija. Trendas yra tam tikra pastovi duomenų savybė, paprastai paaiškinama tam tikrais fiziniais paviršiaus ypatumais. Daugeliu atvejų duomenys neturi trendo arba jis labai silpnas, todėl jį galima ignoruoti. Tokiu atveju priimame, kad duomenys turi pastovią vidutinę reikšmę:  $\bar{z}(s) = \bar{z}$  visose  $s$ . Regresija tai atsitiktinio dydžio vidurkio priklausomybė nuo kitų dydžių (kelių dydžių) radimas. Globalūs metodai yra laikomi netiksliais, kadangi paviršiaus interpoliavimo procedūros apskaičiuoja vieną funkciją, aprašančią visą modeliuojamą paviršių, o reiškiniai stipriai kintant globali erdvinė statistika duoda klaidingus rezultatus. Kadangi tiriant realų pasaulį nagrinėjami duomenys dažniausiai nėra monotoniški yra naudojami ne globalūs, o lokali metodai.

Lokali polinom interpoliacija veikia atvirkščiai – vertina paviršių pagal vienas po kito esančius tinklo susikirtimo taškus, skaičiavimams naudojant tik artimiausius taškų reikšmes. Paviršiaus interpoliavimas susideda iš keturių etapų:

- 1) teritorijos apie interpoliuojamą tašką dydžio radimas;
- 2) artimiausių taškų suradimas;
- 3) matematinio modelio pasirinkimas (jis reikalingas apskaičiuoti paviršiaus reikšmei pagal turimus taškus);
- 4) modelio (paviršiaus) reikšmės radimas.

Lokalus metodas naudoja taškus su žinomomis reikšmėmis tam, kad galėtume prognozuoti nežinomą reikšmę, todėl ir yra vadinamas tiksliau. Lokaliems yra priskiriami splainas, krigingas bei atvirkštinio atstumo metodai, kuriuos plačiau aptarsime sekančiose skirsniuose.

Taip pat paviršiaus modeliavimo metodai gali būti suskirstyti dvi grupes pagal skaičiavimams reikalingus modeliavimo parametrus: metodus, kurie duomenų analizei naudoja variogramas (krigingas) bei metodus, kurie naudoja taškų svorius (splainas, svorio metodas). Svorio metodas – tai toks paviršiaus modeliavimo metodas, kai pradiniam taškų aukščiams priskiriami svoriai, kurie yra atvirkščiai proporcingi atstumui iki skaičiuojamo taško. Tai reiškia kuo labiau yra nutolęs taškas, tuo mažesnis svoris yra jam priskiriamas (Johnston and others, 2001). Svorius naudoja deterministiniai interpoliavimo metodai, o stochastiniai reikšmių radimui naudoja variogramas.

Variograma (semivariograma) – parametrinė funkcija, naudojama gretimų taškų aukščių koreliacijai (statistinei priklausomybei) nustatyti. Variograma naudojama rasti optimaliausius erdvinis sprendimus, vertinti objektų padėtį, dydį. Taip pat ji parodo teritorijų ribas, kuriose išmatuoti dydžiai turi apibrėžto diapazono reikšmes. Variogramų naudojimo privalumas – paprastas matematinis realizavimas, o trūkumas tai, kad ji yra nejautri smulkios (reljefo) detalėms (Chao – yi Lang, 1995). Pavyzdžiui, žinant kai kuriuose taškuose žemės paviršiaus aukščių reikšmes, galima

vertinti tarp jų esančius, nežinomus taškus. Vertinant nežinomus taškus galime sudaryti skaitmeninį tolydų reljefo paviršių. Tam reikia modeliuoti semivariogramas, priglodinti geriausi semivariogramos modelį bei nustatyti semivariogramos parametrus. Pagrindiniai semivariogramos parametrai – plotis, grynuolis ir slenkstis. Slenkstis – tai didžiausias aukštis, kur pasiekia semivariogramos kreivė. Slenkstis dažnai sudaro dvi dalys: trūkio pradžioje, vadinamas grynuoliu ir dalinis slenkstis, kuriuos sudėjus ir gaunamas slenkstis. Plotis – tai nuotolis tarp taškų porų, kuriam esant pasiekiamas slenkstis. Šie trys semivariogramos parametrai padeda kuo tiksliau priglodinti geriausi semivariogramos modelį. Taigi pirmame algoritmo realizavimo etape būtina palaikyti maksimaliai teisingą semivariogramos matematinį modelį. Semivariogramai modeliuoti naudojamos įvairios funkcijos – sferinė, eksponentinė, tiesinė, Gauso (Johnston, Ver Hoef ir kt., 2001). Dažniausiai rekomenduojama sferinė lygtis.

Sferinė semivariograma:

$$\chi(d; H) = \begin{cases} H_s \left[ \frac{3\|d\|}{2H_R} - \frac{1}{2} \left( \frac{\|d\|}{H_R} \right)^3 \right], & \text{kai } 0 \leq \|d\| \leq H_R \\ H_s, & \text{kai } H_R < \|d\|. \end{cases} \quad (1.1)$$

Eksponentinė semivariograma:

$$\chi(d; H) = H_s \left[ 1 - \exp\left(-\frac{3\|d\|}{H_R}\right) \right] \quad (1.2)$$

Gauso semivariograma:

$$\chi(d; H) = H_s \left[ 1 - \exp\left(-3 \left( \frac{\|d\|}{H_R} \right)^2\right) \right] \quad (1.3)$$

čia  $\chi$  – semivariogramos funkcija,  $d = (d_x, d_y)^T$ ,

$\|d\|$  – atstumas tarp taškų;

$H_s \geq 0$  – dalinis slenkstis, kuris apskaičiuojamas iš slenkstio atimant grynuolį;

$H_R \geq 0$  – plotis.

Taip pat reikia nepamiršti, jog esant aiškiai išreikštai reiškinio anizotropijai, nagrinjamos variogramos ir interpoliuojamas paviršius turėtų būti atskirai kiekvienai zonai.

Erdvinė priklausomybė (autokoreliacija) gali priklausyti tik nuo atstumo tarp dviejų taškų, tai vadinama izotropija. Tačiau manoma, kad ta pati autokoreliacijos reikšmė gali atsirasti skirtingais atstumais, matuojant skirtingomis kryptimis (esant didesniems atstumams objektai vienomis kryptimis yra panašesni nei kitomis). Ši pusvariogramose ir kovariacijoje pastebima kryptinė taka vadinama anizotropija. Tiriant anizotropiją, svarbu atsižvelgti tai, kad, nustatius kryptinius skirtumus erdvinėje priklausomybėje, juos galima traukti pusvariogramas ar

kovariacijos modelius, kurie turi tokos taikomam geostatistinio interpoliavimo rezultatui. Ištyrus anizotropij bei priglodinus semivariogramos modelius geriausi iš jų padeda nustatyti kvadratinį nuokrypi sumą. Kvadratinio nuokrypi suma (angl. *square of the sum of differences (SS)*) dar vadinama vertikali atstumu nuo taško iki kreivės kvadrat suma ir apskaičiuojama pagal formulę:

$$SS = \sum_i (y_i - \hat{y}_i)^2, \quad (1.4)$$

čia  $y_i$  - stebimos funkcijos reikšmės;

$\hat{y}_i$  - sumodeliuotos funkcijos reikšmės.

Geriausiai tinkančiu yra laikomas tas modelis, kurio kvadratinis nuokrypi suma yra minimaliausia.

Aukščiau trumpai aptarti stochastinis (krigingas) bei deterministinis (splainas, atvirkštinio atstumo metodas) paviršiaus modeliavimo metodai bei jų skirtumai (Kumetainienė, 2006). Sekančiuose skirsniuose aptarsime kiekvieną metodą išsamiai. Pradysime nuo stochastinio interpoliavimo metodo – kringingo ir jo rėšis.

## 1.2 Stochastiniai modeliavimo metodai – kringingas

Kringingas iš pradžių buvo skirtas ir iki šiol tebėra naudojamas kalnakasyboje nustatyti metalo kiekio uolienos luituose. Tokiomis aplinkybėmis kiekvienas uolienos luitas gali būti tiriamas, ir vertinamas metalo kiekis jame (Bernard et al., 2009). Kalnakasys tada gali nuspręsti ar luite yra pakankamai metalo, kad jį būtų verta išgauti ir apdoroti. Aplinkos mokslininkai, o ypač pedologai (dirvožemio mokslo specialistai), kringingą naudoja kiek kitaip, t. y. optimaliai vairi vietose interpoliacijai, sudarant žemėlapius. Pirmieji tokia būdu ordinarų kringingą panaudojo Burgessas ir Websteris (1980a, 1980b) Burgessas ir kt. (1981). Vėliau kringingas buvo naudojamas kininkavimo srityje Mulla (1997), Frogbrookas (1999) ir Frogbrookas ir kt. (1999). Naudojant kringingo metodą interpoliuojama reikšmė priklauso nuo dviejų veiksnių: trendo bei papildomo kintamumo elemento. Pavyzdžiui, jei kylate nuo jūros aukštyn kalnus, turite aiškiai išreikštą altitudinį<sup>1</sup> trendą. Tačiau tikrasis vietovės reljefas svyruoja: kilns aukštyn ar leis žemyn kirsdamas upės slėnius, bus daugybė kalv ir duburčių. Kringingo interpoliavime yra teigiama, kad apie trendą svyruoja autokoreliuota erdvinė atsitiktinė paklaida. „Atsitiktinė“ reiškia tai, kad svyravimai iš anksto nežinomi. Nors šis metodas plačiai naudojamas jis turi trūkumų – skaičiavimo sudėtingumą (Heinrich, 1994). Reikia išspręsti dideles lygtis sistemas kiekvienai tinklinio modelio ląstei, todėl modeliuojant reljefą kringingo metodu, reikia daug daugiau kompiuterio resursų negu interpoliuojant

<sup>1</sup> **Altitud**, pranc. l'altitude – vietovės, statinio, bet kurio kito objekto ar jo taško vertikalus aukštis nuo sutartinio atskaitos taško. Paprastai geografijai, geodezijai, kartografijai, meteorologijai, architektūrai, aviacijai, fizikai, biologijai atskaitos tašku tarnauja jūros lygis.

pavirši kitais reljefo modeliavimo metodais. Daug d mesio bei laiko b tina skirti variogramos formai, duomen analizei bei gretim zon dydži nustatymui. Kai paviršius modeliuojamas kriginogo metodu, rezultatai, kurie yra reikalingi paviršiui sudaryti yra generuojami pagal pradini duomen reikšmes.

Bet kokio uždavinio sprendimas kriginogo metodu randamas vadovaujantis stochastin s erdvin s variacijos tolydžiuoju modeliu. Jis geriausiai panaudoja turim informacij , atsižvelgdamas dydžio kitim erdv je pagal variogramos model . Variograma yra efektyvus rankis prognozuoti neišmatuot tašk reikšmes su išmatuotais taškais, naudojant atstum ir krypt bei atrinkti tokios informacijos klasterius<sup>2</sup>. Iš pradži kriginogas buvo suprantamas, kaip papras iausia tiesin aplinkini duomen svertini vidurki suma. Po to kriginogo metodas buvo pakeistas ir pritaikytas spr sti daug sud tingesnes užduotis kalnakasyboje, naftos inžinerijoje, užterštumo kontrolei ir mažinimui bei visuomen s sveikatos srityje. Šiuo metu terminas „kriginogas“ yra bendrinis ir apima kelet skirting kriginogo tip - tiek tiesini (su trendo modeliu), tiek netiesini , kurie skirti darbui su tr kiais arba dvejetainiais duomenimis. Iš pradži aptarsime paprastesnius, tiesinius metodus. Tiesiniame kriginoge ver iai yra svertin s tiesin s duomen kombinacijos. Imties duomenims esantiems aplink tašk ar blok priskiriamos reikšm s, kad kriginogo variacijos vertis b t minimalus. Šiuo poži riu kriginogas yra optimalus. Pagal paskirt kriginagai skirstomi kelias r šis (John Wiley et al., 2007).

Normalusis arba ordinarus kriginogas yra grindžiamas nežinomu, ta iau pastoviu  $\sim$  , ir laikoma, kad autokoreliuotos erdv je duomen reikšm s (atsitiktin s paklaidos) svyruoja apie j . Kadangi tikslaus  $\sim$  žinoti praktiškai ne manoma reikia žinoti tai, kad fiziškai pagr stas modelis duoda trend . Tada galima imti skirtum tarp modelio ir liekan , ir naudojant paprast j kriginog vertinti liekanas. Apskai iavus trend jos turi b ti lygios 0 (Webster, 2007).

Paprastasis kriginogas tai papras iausias interpoliavimo metodas. Jis yra naudojamas, kai yra žinomas vidurkis (vidurkio funkcija). Jis praktiniuose uždaviniuose taikomas retai, nes nelabai tik tina, jog vidurkio funkcija bus žinoma. Tod l paprastasis kriginogas pakei iamas bei pritaikomas kitomis formomis, kuriuose duomenys transformuojami taip, kad b t žinomi vidurkiai. Po duomen transformacijos gali b ti naudojami indikatorinis arba disjunkcinis kriginogas.

Disjunkcinis kriginogas taip pat yra netiesinis kriginogo metodas, bet jis yra griežtai parametrinis. Priimant sprendimus nustatomi ne tik ver iai bet ir tikimyb viršyti ar neviršyti nustatyt ribin vert . Viena iš svarbiausi disjunkcinio kriginogo savybi ta, kad jis tinka dirbant su tr kiais arba dvejetainiais duomenimis. Šia savybe pasižymi ne tik disjunkcinis, bet ir kitas kriginogo metodas – indikatorinis.

---

<sup>2</sup> Klasteris (angl. cluster) — vienetas, sudarytas iš keli vieno tipo element , kuris veikia kaip vieninga sistema, turinti tam tikras savybes.

Indikatorinis krigingas yra netiesin, neparametrin krigingo forma, kurioje tolydieji kintamieji pakeičiami dvinariais (indikatoriais). Šis metodas populiarus, nes juo galima vertinti beveik visus skirstinius<sup>3</sup> ir apskaičiuoti vertinimą empirini skirstiniais, taip nustatant jų patikimumo ribas. Jis taip pat galima traukti kokybinę informaciją, kad prognozė būtų tikslesnė. Indikatoriaus koncepciją galima pritaikyti tolydžiajam skirstiniui, kurio metu duomenys transformuojami nulius arba vienetus, priklausomai nuo (a) vertės pozicijos nustatytos ribinės vertės atžvilgiu arba (b) geologinio požymio buvimo ar nebuvimo. Tarkime, jeigu analizuojamoje vietoje nėra mineralinių medžiagų, jai galima priskirti nulį, o jei jos yra – vieneto vertę. Tokiu pat principu, viršijantiems sutartin ribinių reikšmių (didžiausi ar mažiausi galimi reikšmės), galima priskirti vienetą, o esantiems žemiau ribinės reikšmės – nulį. Žinoma, naujai transformuotam kintamajam reikia nustatyti semivariogramos modelį, ir tada galima atlikti krigingą kaip nurodyta anksčiau. Krigingo su indikatoriniais vertinimais rezultatas skirtas vertinti santykiu  $p_k$ , kur galima aiškinti dviem būdais (Alastair, 2002):

- (a) Tikimybė, kad matavimo vietoje reikšmės yra priskiriamas vienetą lygi  $p_k$ ;
- (b) Jeigu vienetą yra priskiriamas taškams aplink išmatuotą teritoriją, reiškia egzistuoja reikšmės, kurioms priskiriamas nulis ir jų tikimybė  $(1 - p_k)$ .

Tuomet, kai indikatorinis krigingas kartojamas skirtingoms ribinėms reikšmėms, procesas vadinamas kartotiniu indikatoriniu krigingu ir tai leidžia vertinti sumines kreives, iš kurios galima nustatyti vietinį vidurkį, vertinti blok skirstinį ir blok, esantį virš ribinės reikšmės, santykiu. Tam, kad gautos reikšmės būtų dar tikslesnės indikatorinis krigingas buvo patobulintas ir pavadintas tikimybinis.

Tikimybinis krigingas buvo pasiūlytas Sullivano 1984 m., o vėliau aprašytas bei išanalizuotas Chileso (1999) bei Goovaertso (1997). Šis krigingo tipas yra pranašesnis už indikatorinį, kadangi atsižvelgiama į tai, kiek reikšmė yra arti ribinės. Tam, kad vertinti indikatorinį kokrigingo metu yra naudojamos pagal dydį išdėstytos  $z(x)$  reikšmės, kurių suma būtų lygi vienetui. Kokrigingas – tai interpoliacinis būdas, leidžiantis tiksliau apskaičiuoti kelis kintamuosius, kurie yra priklausomi vienas nuo kito. Šis metodas padeda pagerinti analizuojamos vietos vertinimą, naudojant susijusius kintamuosius, kurie yra žinomi ne tik tuose vietose. Taip pat šis metodas naudojamas siekiant užtikrinti vertint reikšmių sąsają, kai kintamieji yra funkciškai susiję. Petitgasas (1991 m.) pasiūlytą kokrigingo modeliu vertinti žuvų susibūrimą bei išsibarstymą tvenkinio teritorijoje dienos ir nakties metu. Žuvų susibūrimas dieną ir naktį buvo susijęs, kadangi naktį žuvų tankis tam tikrame taške lygus tankiui kažkuriame plote dienos viduryje. Ryšys leidžia išskirti variogramas ir išvesti

<sup>3</sup> Skirstinys (tikimybinis pasiskirstymas ar pasiskirstymo dėsnis; angl. probability distribution) – tai požymio reikšmės, arba atsitiktiniai dydžiai, ir jų tikimybi tarpusavio ryšys. Požymio empiriniai skirstiniai *pagrindinė charakteristika yra aritmetinis vidurkis*. Jam būdinga tai, kad nuo jo teigiamai ir neigiamai nuokrypių suma lygi nuliui.

santykini s variogramos model tarp dienos ir nakties verti . Kokrigingo proced ra leidžia vertinti dienos (arba nakties) žem lap , naudojant ir dienos, ir nakties pavyzdžius.

Ordinarusis kokrigingas yra ordinaraus vieno kintamojo krigingo pl tinys dviems ar daugiau kintam j . Kad jis b t efektyvus, tarp kintam j turi b ti tam tikra koregionalizacija<sup>4</sup>. Jis itin naudingas, jei koks nors dydis, kur galima lengvai išmatuoti, yra erdviškai susij s su kitu ar kitais dydžiais, kuriuos išmatuoti yra brangu, ir jie mažai kur matuojami. Kokrigingas leidžia tiksliau vertinti mažesn s imties dydžius, naudojant dažniau matuojam dydži erdvin informacij .

Tiek ordinar , tiek skyriaus pradžioje min t paprast j kriging galima atlikti su transformuotais duomenimis. Taigi, lognormalusis krigingas yra tiesiog logtransformuot duomen krigingas – procesas analogiškas neapdorot duomen krigingui, naudojant santykin semivariogram . Ši semivariogram galima vadinti nestacionarumo koregavimo priemone, tuo atveju, jei vietos dispersija yra proporcinga vietos vidurkio kvadratui. Ji naudojama labai asimetriškiems duomenims, kurie art ja prie lognormaliojo skirstinio.

Universalusis krigingas yra tiesinis su nežinomu polinominiu trendu. Jis yra mišrus interpoliavimo metodas, kurio sistemoje globalus trendas modeliuojamas kaip geografini koordinata i funkcija. Tai krigingo proced r papildo didžiausio tik tinumo metodu. Universaliajame krigingo modelyje trendas atskiriamas nuo liekam j duomen ir modeliuojamas atskirai. Matheronas (1971) pademonstravo, kad trendas ir liekamieji duomenys negali b ti vertinami kartu naudojant tik vien atsitiktin s funkcijos (t. y. vien duomen aib iš 1 met ) realizacij : ( vertint ) likutini reikšmi variacija po erdvinio išlyginimo proced ros vertinant trend yra mažesn nei (tikroji) likutin proceso variacija. Rivoirardas ir Guiblinas (1997) si lo atsižvelgti paklaid , atsirandan i tiriant trend , kuri reikalinga vidurkio ver io variacijos vertinimui. Pagalbiniai kintamieji naudojami vertinti trend regresijos metodu. Naudojant š metod kintamieji turi b ti tarpusavyje nekoreliuoti (statistiškai nepriklausomi).

Dar vienas krigingo metodas, kuris gali b ti pritaikomas duomenims, kur yra trendas – medianinio išlyginimo krigingas. Medianinio išlyginimo proced ros metu iš pradini duomen yra pašalinamas trendas, teigiant kad liekanos yra stacionarios (vidurkis pastovus, bet nežinomas). Liekanoms parenkamas bei priglodinamas vienas iš parametrini semivariogramos modeli (Gauso, eksponentinis arba sferinis) bei taškiniams ar bloko duomenims pritaikomas ordinarus krigingas. Taip atsitiktin funkcija yra suskaidoma deterministin trend ir atsitiktin liekan komponent . Išprognozavus liekanas jos yra sudedamos su eilu i , stulpeli ir bendrais median efektais. Tiau jei apie trend n ra jokios informacijos tuomet gali b ti naudojamas Bajeso krigingas.

---

<sup>4</sup>Koregionalizacija - regionini kintam j krigingo teorijos išpl timas iki keli kintam j , turin i daugiamat kryžmin erdvin koreliacij bei atskiras vienmates erdvines autokoreliacijas .

Bajeso kriging pasi l Omre (1987) tokiems atvejams, kai duomenyse yra trendas. Tai tarpinis variantas tarp paprastojo krigingo, kuris naudojamas, kai nėra trendo, ir universaliojo krigingo, kuris naudojamas, kai jis yra žinomas. Krigingo lygtys yra tokios pat kaip paprastojo krigingo, bet su nestacionariomis kovariacijomis (Chilesas ir Delfineris, 1999).

Atlikus vien iš išvardint kriging gali atsirasti neigiamos reikšmės, kuri paprastai galima išvengti arba ją poveik sumažinti tam tikromis priemonėmis:

a) užtikrinant, kad duomenis nedengia kiti duomenys, esantys arčiau vertinamo taško;

b) naudojant krigingo metodą;

c) patikrinant, ar neatsirado neigiamos reikšmės po krigingo, atmetant duomenis su neigiama verte ir iš naujo atliekant likusių duomenų krigingą;

d) kaip atsargumo priemonę, nutolusias vertes vertinant atskirai – net ir nedidelė neigiama vertė pritaikyta nutolusioms vertėms gali duoti kraštutinius rezultatus ir kai kuriais atvejais, netgi neigiamas laipsnio reikšmes.

Jeigu dėl sudėtingų skaičiavimų krigingo metodas atmetamas, tada specialistai pasirenka kitą erdvinio prognozavimo būdą. Greta vairių rimtų mokslinių literaturoje siūlomų alternatyvų (pvz. nuosekliojo interpoliatoriaus ir neparimetrinio erdvinio regresijos metodo) neretai yra siūlomi deterministiniai metodai.

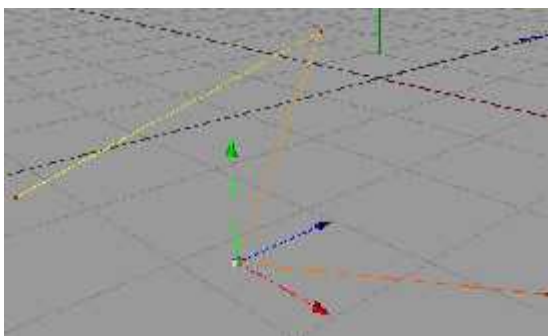
### 1.3 Deterministiniai metodai

Praktikoje dažniausiai naudojami deterministiniai interpoliavimo metodai - splainas ir atvirkštinio atstumo metodas. Vienas iš populiariausių ir dažniausiai taikomų deterministinių metodų yra interpoliacija splainais. Terminas splainas (angl. *spline*) - reiškia prietaiso pavadinimą, kurį naudojo braižymui, tam kad tolygiai sujungti taškus (Albert J., Nilson E., 1972). Naudojant šį metodą taškai kuriuose jungiasi kreivės segmentai yra vadinami mazgais. Splaino metodo pagrindinė savybė yra, kad duotojo mazgo aplinkoje interpoliacijai yra naudojamos netolydžios polinominės funkcijos, kurios taip pat vadinamos splaino funkcijomis. Šis metodas glaudžiai susijęs su krigingu, be to vairūs empiriniai palyginimai leidžia teigti, kad šio metodo prognozinių charakteristikų dažnai panašios (Hutchinson ir Gessler 1994; Laslett ir McBratney 1990; Laslett, McBratney, Pahl ir Hutchinson 1987; Voltz ir Webster 1990). Atitinkami splaino ir krigingo privalumai buvo aptarti žurnalo „Amerikos statistikas“ (angl. *The American Statistician*) 1990 m. rugpjūčio mėnesio numerio laiškuose redaktoriui. Wahba (1990a) suabejojo Cressie's (1989) teiginiu, kad „krigingas, kitaip negu splainai, pritaikomas, atsižvelgiant duomenų atskleidžiamos erdvinės priklausomybės kokybę ir kiekybę ir prieštaravo teigdamas, kad „vairūs teiginiai, jog krigingo būdu gauti vertinimai kažkodėl geresni negu splaino metodu gauti vertinimai, neturi prasmės“. Splaino metodas gali būti pasirinktas kaip alternatyva krigingui iš dalies todėl, kad jis

atlieka deterministinio erdvinio prognozavimo metodo vaidmenį, ir iš dalies todėl, kad atsako klausimus kylančius iš esamo šio dviejų metodų palyginimo.

Splainas – tai tolydžioji iki  $p$ -tosios eilės (imtinai) išvestinė funkcija, sudaryta iš kurios nors funkcijos dalių. Splainais vadinamos linijos jungiančios mazgus ir išsilenkiančios tarp jų tam tikru nustatytu laipsniu. Splainą sudaro du pagrindiniai elementai: viršūnė ir segmentas. Viršūnė tai taškas su tam tikromis XYZ koordinatėmis erdvėje, o segmentas tai linija jungianti dvi artimas viršūnes. Istoriskai splainai pradėti konstruoti iš  $n$ -tosios eilės polinomo dalių, todėl splainus galime laikyti funkcijomis, sudarytomis iš polinomo dalių. Funkcijos, kurias dabar vadina splainais matematikams žinomos jau seniai, nuo Eilerio laikų, tačiau jų intensyvus analizavimas bei tyrimas prasidėjo tik XX amžiaus viduryje. 1946 metais Isaakas Schenbergas pirmasis panaudojo šį terminą kaip polinominių splainų klasės apibrėžimą. Iki 1960 metų splainai buvo pagrindinė teorinė tyrimų priemonė, jie dažnai buvo naudojami sprendžiant variacines ekstremalias bei variacines užduotis. Dabar splainai yra naudojami tiek matematikos teorijoje, tiek ir vairiuose skaitiniuose taikymuose. Splainai turintys du kintamuosius yra intensyviai naudojami užduotims, susijusioms su paviršiumi, atlikimui. Tai yra daroma naudojantis variomis kompiuterinio modeliavimo sistemomis (Wikipedia, 2010). Splainai kaip ir kringas pagal savo charakteristikas yra skirstomi kelias rūšis. Rūšis yra nustatoma remiantis segmentų jungiančių viršūnės savybėmis. Segmentai gali būti varių tipai: tiesūs arba išlenkti, su aštriais kampais arba interpoliuoti, jie gali kirsti viršūnes arba būti šalia jų. Priklausomai nuo segmento tipo yra išskiriamos splainų rūšys: paprastasis (tiesinis), kubinis, Akima, B-splainas ir Bezier ir t.t.

Paprastasis splainas – toks, kuris yra sudarytas iš pirmojo laipsnio polinomo. Interpoliacijos tikslumas tiesiniais splainais nėra aukštas, kadangi segmentai yra jungiami tiesiomis linijomis ir prie viršūnių susidaro aštrūs kampai. Taip pat reikia paminėti, kad jie neįtikrina net pirmą išvestinį tolydumą. Tačiau kai kuriais atvejais tiesinė aproksimacija su trūkiais taškais gali būti naudingesnė, nei aukštesnės eilės aproksimacija. Vienas iš privalumų yra tai, kad tiesinis splainas išsaugo perduotą taškų monotoniškumą.

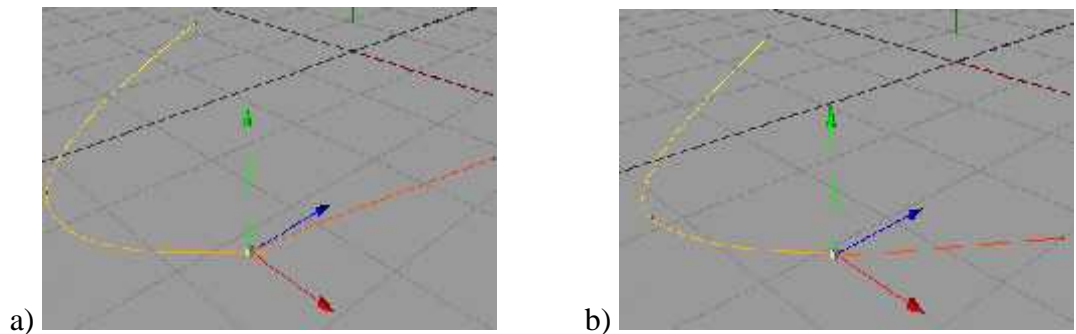


1.1 pav. Aštrūs kampai tiesinio splaino viršūnėse

Ermito splainas – tai kubinis splainas (treios eilės), kurio išvestinėgyja reikšmes splaino mazguose. Kiekviename Ermito splaino mazge yra užduota ne tik funkcijos reikšmė, bet ir

pirmosios išvestinės reikšmės. Šis splainas turi tolydžią pirmąją išvestinę, bet antroji išvestinė jau turi trūkio tašką. Interpoliacijos tikslumas žymiai geresnis nei tiesinio splaino.

Kubinis splainas – tai toks splainas, kuris yra gaunamas jei pirmoji ir antroji išvestinės yra tolydžios (ALGLIB, 2010). Jo segmentai jungiantis virš nes yra kreivės, kurios gali turėti nuokrypį, todėl šis nuokrypių splainas atrodo tolygus (1.2 pav.).



1.2 pav. a) Kubinio splaino kreivė, b) akimos splaino kreivė

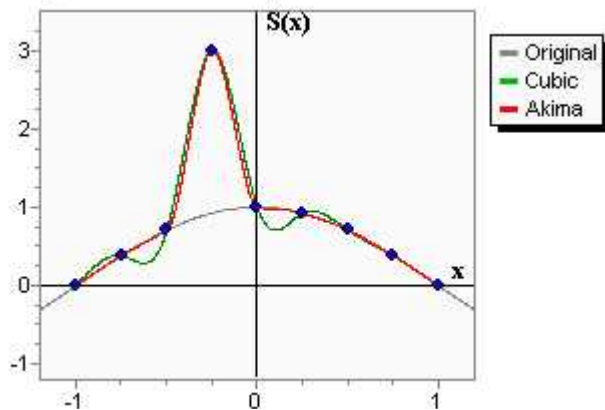
Kubinis splainas yra suformuojamas funkcijos reikšmėmis mazguose ir išvestinė reikšmėmis interpoliacijos intervalo galuose (1.2 pav. a)). Jeigu yra žinoma tiksli pirmosios išvestinės reikšmė abiejuose intervalo galuose, tuomet toks splainas yra vadinamas fundamentaliuoju. Interpoliacijos paklaida naudojant šį splainą yra lygi  $O(h^4)$ . Čia  $h$  yra vadinamas žingsniu, kuo jis mažesnis tuo tikslesni yra matavimai. Jeigu pirmosios (arba antrosios) išvestinės reikšmės interpoliacijos intervalo galuose yra nežinomos, tuomet galima užduoti kraštines sąlygas  $S''(A)=0$ ,  $S''(B)=0$ , tuomet turime natūralų splainą. Interpoliacijos paklaida naudojant šį splainą lygi  $O(h^2)$ . Maksimalios paklaidos yra pastebimos ribiniuose mazguose, vidiniuose mazguose interpoliacijos tikslumas yra žymiai didesnis. Jeigu yra nežinomos ribinės funkcijos išvestinės reikšmės tuomet galima panaudoti dar vieną kraštinių sąlygą: „splainas, kuris baigiasi parabolė“. Šiuo atveju ribinio splaino atkarpa yra išreiškiama antrojo laipsnio polinomu vietoj trečiojo (vidiniams atkarpoms naudojamas trečiojo laipsnio polinomas). Šios sąlygos išpildymas daugeliu atveju garantuoja didesnį tikslumą, nei kraštinių sąlygos. Pagaliau galima derinti vairių tipų kraštinių sąlygų viariose ribose. Paprastai taip yra daroma tuomet, kai turime tik dalį informacijos apie funkcijos elgesį ribiniuose taškuose. Tačiau ir visų sąlygų išpildymas duos mažiau tikslūs rezultatus nei interpoliacijomis Akimos splainais.

Akimos splainas – tai ypatingas splaino tipas, kuris yra atsparus išskirtims. Kubinio splaino trūkumas, kad jie link osciliuoti<sup>5</sup> taško, kuris žymiai skiriasi nuo gretimų taškų, aplinkoje, o Akimos splainai neturi nuokrypių ir interpoliuota kreivė eina tiksliai per virš nes. Dėl šios priežasties ji yra tikslesnė bei staigesnė (1.2 pav. b)).

Svarbi Akimos splaino savybė yra jo lokalumas – funkcijos reikšmės intervale  $[x_i, x_{i+1}]$ , priklauso tik nuo  $f_{i-2}, f_{i-1}, f_i, f_{i+1}, f_{i+2}, f_{i+3}$ . (ALGLIB, 2010). Antroji savybė, kuri taip pat nra

<sup>5</sup> osciliacija [lot. oscillatio — svyravimas], per lygius laiko tarpus pasikartojantis svyravimas, virpesys.

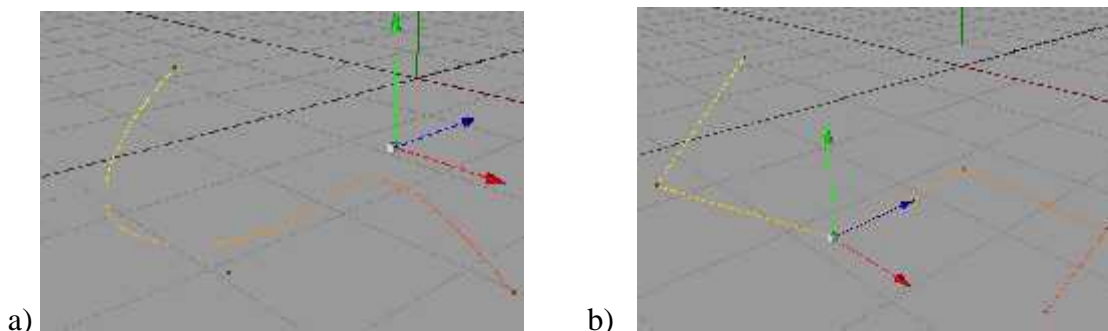
mažiau svarbi, yra interpoliacijos Akimos splainais netiesiškumas t.y. dviej funkcij interpoliacijos suma n ra lygi interpoliacini schem , sudaryt remiantis atskiromis funkcijomis, sumai. Tam, kad sudaryti Akimos splain reikia tur ti ne mažiau 5 tašk . Vidin je srityje interpoliacijos paklaida turi eil  $O(h^2)$ .



1.3 pav. Akima ir kubinio splaino grafinis palyginimas

Paveiksle (1.3) yra pavaizduotas tašk rinkinys, kuriame vienas taškas stipriai skiriasi nuo savo gretim reikšmi . Žalia spalva yra pažym tas kubinis splainas su kraštin mis s lygomis. Nesunku pasteb ti, jog interpoliacijos intervaluose, kurie ribojasi su nuo kit nutolusiu tašku, splainas nukrypsta nuo interpoliuojamos funkcijos – tai yra išskir i pasekm . Raudona spalva yra pažym tas Akimos splainas. Galime pasteb ti, kad š splain praktiškai ne takoja trendas. Intervaluose, kurie ribojasi su nutolusiu tašku beveik n ra osciliacijos (svyravimo) požymi . Šiuo ir prieš tai aprašyta is atvejais splain segmentai b tinai turi kirsti mazgus (virš nes), ta iau sekan iu atveju yra kitaip.

Interpoliuojant B – splainais kreiv eina ne per visas virš nes, o tik linksta j kryptimi. Kuo virš n yra ar iau kreiv s tuo didesn tak jai ji turi.



1.4 pav. a) B – splain kreiv , b) Bezier kreiv

B - splainas yra sudarytas iš keli Bezier<sup>6</sup> segment su tam tikru t stinumo laipsniu, lokalizuojant kontrolini tašk poveik kreivei bei pridedant svorius. Savyb s: visi taškai, esantys

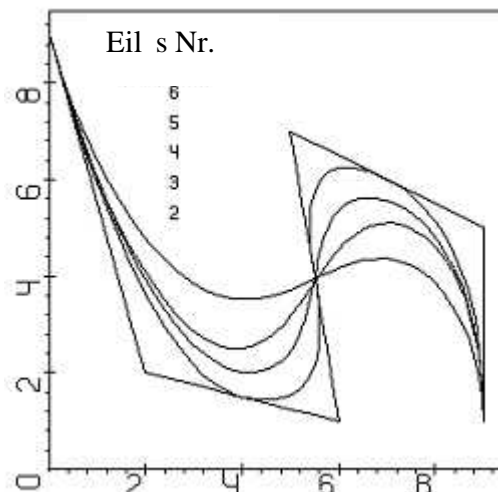
<sup>6</sup> Bezier kreiv s – kompiuterin je grafikoje naudojamos glotnios parametrizuotos kreiv s. Bezier kreives nesunku nupiešti turint tik kelet parametrizuojan i tašk . Daugiamat je erdv je Bezier kreives pakei ia Bezier paviršiai.

tarpo 1 ir  $n$ -tojo atraminio taško, yra interpoliuojami, kreivės liestinė 1 atraminiam taške yra ant atkarpos, jungiančios 1 ir 2 atraminius taškus; kreivės liestinė  $n$ -tame atraminiam taške yra ant atkarpos, jungiančios  $n$  ir  $n-1$  atraminius taškus. Bezier kreivė yra išsidėsčiusi iškilame plote, ribojamame atraminio tašku (VDU, 2009). Bezier kreivės kreivum galimi reguliuoti. Keičiant viršūnių ilgį, galima reguliuoti segmento kreivumą, o pakeitus kryptį nustatyti kreivei liestinę tame taške t.y kreivumo kryptį (1.4 pav). Šio tipo kreivės naudojamos dėl patogumo skaičiuoti kompiuteriu, stabilumo slankaus kabelio klaidų atžvilgiu, mažo atminties suvartojimo, reikalingo konstrukcijos saugojimui ir galimybių gyti bet kokią formą. Bezier metodo pagrindą sudaro aproksimacija<sup>7</sup> Bernsteinio daugianariais<sup>8</sup>. Tai galima pasirinkti ir kitokius pagrindinius funkcijų rinkinius. Jei lyginsime Bernsteinio daugianarius ir B – splineus, tai pastarieji turi tokias ypatybes:

a) Aproksimacija B – splineais garantuoja tikslesnį laukto artumą prie analizuojamos reikšmės, nei aproksimacija Bernsteinio daugianariais.

b) Kai yra aproksimuojama Bezier metodu, tuomet kiekvienai laukto koordinačiai reikšmei turi būti taikomi visi Bezier kreivės viršūnės. Dėl šios priežasties sunkiau koreguoti atskirus kreivės plotus. O aproksimacija B – splineais pasižymi lokališkumu, todėl galima padaryti lokalius kreivės pakeitimus be visų reikšmių perskaičiavimo.

c) Vienintelis Bezier būdas padidinti (sumažinti) kreivės eilę yra viršūnių skaičiaus bei atitinkamos Bezier laukto padidinimas (sumažinimas). B – splineo metode šie du parametrai yra nepriklausomi, todėl gali būti išrenkamas bet kuris.



1.5 pav. B – splineai turintys skirtingą eilę

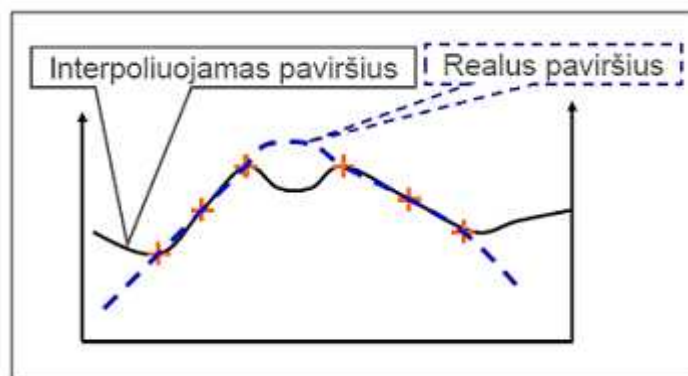
Paveikslėlyje (1.5) pavaizduotos, kai kurių eilių B – splineo kreivės. Iš paveikslėlio galime pastebėti, kad augant eilei, B – splineo kreivės darosi tiesesnes ir jų forma vis labiau skiriasi nuo aproksimuojamos laukto.

<sup>7</sup> Aproksimacija – tai radimo procesas tokios kreivės, kuri geriausiai atspindėtų turimų duomenų kitimo pobūdį.

<sup>8</sup> Bernsteinio daugianariai – tai algebriniai daugianariai, sudarantys tiesinį Bernsteinio daugianarių kombinaciją.

Multivariate Adaptive Regression Splines (angl. *MAR Splines*) – daugialypiai adaptyv s regresiniai splainai (2001; Multivariate Adaptive Regression Splines, *Annals of Statistics*, 19, 1-141). Šie splainai paskutiniu metu labai dažnai naudojami sudarant sudingus prognozavimo modelius, kai nėra pastebima paprasta ar monotoniška ryšio tarp nepriklausomų ir priklausomų kintamųjų. Dėl šiuo metodu parenkam nepriklausomų kintamųjų specifiškumo MAR - splainai tinkami kaip patogus naudoti regresijos medžiūs. Iš tikrųjų, vietoj to, kad laikyti šio apibendrintu daugialyp s regresijos metodu MAR – splainai gali būti nagrinjami kaip regresijos medžiū apibendrinimas, kuriuose binarinis pasiskirstymas pakeistas tolydžiomis bazinėmis (pradinėmis) funkcijomis. Daugialypiai adaptyv s splainai (MAR - splainai) tai yra Friedmano metod apibendrinimas skirtas regresini ir klasifikacijos uždaviniams sprendimui. Regresijos uždaviniai yra sudaryti ryšio tarp daugelio priklausomų kintamųjų (dar vadinam atsako), vieno ar kelių nepriklausomų kintamųjų (vesties kintamųjų arba prediktorių) nustatymui. Priklausomas kintamasis – kintamasis, kurio reikšmės mes norime nuspėti naudojant nepriklausomo kintamojo reikšmę. Tokie splainai, kuriuose reikia nuspėti priklausomų kintamųjų reikšmes turint nepriklausomus kintamuosius yra skirti kategorini bei tolydži kintamųjų apdorojimui nepriklausomai nuo to ar jie yra priklausomi kintamieji ar nepriklausomi. Jeigu tai yra kategoriniai kintamieji, tuomet MAP – splainai sprendžia klasifikavimo uždavinį. Priešingu atveju, jeigu priklausomi kintamieji yra tolydūs tuomet uždavinys yra sprendžiamas kaip regresinis. Dažniausiai programose naudojant MAP – splainus uždavinio tipas yra parenkamas automatiškai.

Dar vienas dažnai naudojamas deterministinis metodas – atvirkščiai proporcingo atstumo (angl. *inverse distance weighted* (IDW)). Šia duomenų reikšmėmis nustatomos kaip vidutinis svertinis dydis pagal žinomas reikšmes, kaip svertis naudojant atstumą tarp taškų. IDW metodo panaudojimo būtiną sąlyga yra, kad interpoliuojamam polinom reikšmių intervalas negali skirtis nuo minimalios ar maksimalios žinomų reikšmių.



**1.6 pav. IDW metodo panaudojimo kriterijus**

Taip pat IDW apskaičiuoja gardelių reikšmes, naudodamas tiesinių taškų svorių kombinaciją. Ši technika vertina reikšmę nežinomame taške, artimesniems taškams priskirdama didesnį svorį.

Taip sukuriamas atvirkštinis ryšys tarp svorio ir atstumo. Pagrindinis IDW metodo savybės: metodas gali suteikti gerą preliminarų interpoliuojamo paviršiaus aprašą; duomenims nėra reikalingos prielaidos ir nėra prognozavimo paklaidų vertinimo; IDW veikia geriausiai tuomet, kai atrankos taškai yra tankūs bei tolygiai išdėstyti.

#### 1.4 Krigingas ir splain charakteristikų palyginimas

Erdvinis prognozavimas naudojant negausius matavimus yra ypatingai svarbus daugelyje mokslinių sričių. Šiame poskyryje lyginami du plačiai naudojami metodai – splain ir krigingo metodai, – taikyti keliems realių duomenų rinkiniams.

Krigingo metodas iš pradžių buvo vertinimas palankiai, vėliau atsirado vis daugiau mokslininkų pareiškiančių savo kritiką: nuo švelnaus skeptiško vertinimo iki visiškai neigimo (Akima, 1975; Philip, 1986; Sibson, 1981; Sichel, 1966; Whitten, 1966). Svarbiausi prieštaravimai krigingo metodui rima si tuo, kad šis metodas grindžiamas atsitiktiniais procesais bei stacionarumo naudojimu (Deutsch and Journel, 1992, p. 67). Yakowitz ir Szidarovsky (1985) pristatė modelius, kuriuose neparametrinis regresijos būdu prognozuojama beveik taip pat, kaip ir taikant kriging koreliuojantiems duomenims, ir geriau nei pirmuoju atveju. Buvo padaryta išvada, kad neparametrinė regresija yra patikimesnė. Hutchinson ir Gessler (1994) teigė, kad ir splainai, ir krigingas leidžia padalyti stebimus signalus ir triukšmą, ir atskleidžia glaudų ryšį tarp erdvinio koherentinio (sinchroniško) signalo, modeliuojamo splainais, ir erdvinio koreliuotojo krigingo signalo. Jie pateikė nuorodas keliems duomenų rinkiniams, kuriems splain ir krigingo metodais gauta prognozė buvo beveik tokia pat gera. Azari ir Muller (1992) taikė savo pačių sudarytą paviršiaus išlyginimo funkciją dviem duomenų rinkiniams ir rodė, kad ji bazinį modelį atskleidžia geriau negu krigingas. Wahba (1990) teigė, kad itin koreliuoti triukšmą gali būti sunku atskirti nuo signalo. Autoriai teigė, jog vienas iš esminių krigingo metodo trūkumų tai, kad metodai, grįsti pastoviais atsitiktiniais procesais, negali būti tinkamai pritaikyti kai kurių tipų erdvinis struktūroms. Plačiau apie splaino bei krigingo skirtumus, naudojant eksperimentinius duomenis, buvo aptarta Geoffrey straipsnyje.

Straipsnyje (Geoffrey, 1994) pristatomas sisteminis splain ir krigingo metodai, taikyti realiems duomenims, palyginimas, siekiant nustatyti aplinkybes, kurioms esant, krigingas yra geresnis prognozavimo metodas, ir paaiškinti, kodėl. Siekiant paprastumo, buvo naudojami vienmačiai erdviniai duomenų rinkiniai. Vienas kontrolinis duomenų rinkinys, sudarytas iš praktiškai nekoreliuojančių stebimų. Visi kiti duomenų rinkiniai yra to paties tipo, kuriam geostatistikas gali naudoti modelį, sudarytą iš stacionaraus atsitiktinio proceso su vyraujančiomis teigiamomis nedidelio intervalo autokoreliacijomis (erdvinis klasikinis koreliacijos atitikmuo).

Jeigu šiomis aplinkybėmis krigingo metodas negali veikti geriau negu spline, tuomet jį galima atmesti. Šie du metodai buvo palyginti ištrynus duomenis, o tada juos prognozuojant. Nustatyta, kad kai kuriais atvejais krigingo metodas buvo kur kas tinkamesnis negu spline. Vairios duomenų pašalinimo konfigūracijos atskleidė, kad duomenų charakteristikos ir imties tvarka lemia, kokiais atvejais krigingas yra pranašesnis už spline metodą: jeigu imties tvarka yra „tinkama“, tada ir krigingo, ir spline metodai yra pakankamai tinkami prognozuoti net ir tokius duomenimis, kurių koreliacijos koeficientas parodo stiprų ryšį tarp duomenų.

Iš Geoffrey straipsnio (Geoffrey, 1994), kur yra atliekamas spline ir krigingo empirinis charakteristikų palyginimas kyla du praktiniai pastebėjimai. Pirmasis duomenų rinkiniams, kurie (pritaikius variuos diagnostavimo metodus – Bajes ir t.t.) buvo sudaryti pastovaus atsitiktinio proceso būdu su autokovariacija ir monotoniškai mažėjanti iki 0 didėjanti atsilikimui. Antrasis pastebėjimas, jog krigingas, kaip prognozavimo metodas, gali būti pranašesnis už aproksimaciją splineais pagal vieno arba dviejų veiksnių vidutinį kvadratinį prognozes paklaidą, jeigu: a) duomenų išdėstymas yra reguliarus kvadratinis, autokoreliacijos funkcija pradžioje yra palyginus nedaug pertraukiama, o autokoreliacija tarp gretimų duomenų reikšmi yra teigiama ir silpna arba b) duomenys išdėstyti netaisyklingai, greta esanti duomenų autokoreliacija aukšta, o prognozė reikia atlikti tarp gerai išskirtų duomenų vietų (pvz. vietų, kurių autokoreliacija artima 0).

## II. STOCHASTINI BEI DETERMINISTINI METOD PANAUDOJIMO MATEMATIN INTERPRETACIJA

### 2.1 Splain funkcijos bei j panaudojimas

Šiame skirsnyje bus pateikti pagrindiniai splain teorijos teiginiai, kurie yra reikalingi splainini aproksimacij tyrimui. Tai yra funkcijos glodinimo uždaviniai, kuriuose aproksimacija fig ruoja kaip tarpinis tyrimo etapas (skaitinis diferencijavimas ir integravimas).

Tipinis glodinimo uždavinys yra interpoliacija. Interpoliuojant pagal užduot skai i lentel  $(x_i, f(x_i)), i = 0, \dots, N$ , su tam tikra paklaida, atkarpoje  $[a, b]$  atstatoma funkcija  $f(x)$  ( . . ., 1990). Klasikinis sprendimo metodas sudaromas iš Lagrandžo interpoliacinio daugianario, kuris užduodamas lygtimi:

$$L_N(x) = \sum_{i=0}^N f(x_i) \frac{\tilde{S}_N(x)}{(x-x_i)\tilde{S}'_N(x_i)}, \quad x_i, i = 1, \dots, N, x_i \in [a, b] \quad \tilde{S}_N(x) = \prod_{i=0}^N (x-x_i). \quad (2.1)$$

Nors pagal Vejerštraso teorema, bet kokia netolydi funkcija  $f(x)$  atkarpoje  $[a, b]$  gali b ti kaip manoma priartinta prie daugianario, praktinis Lagrandžo daugianario panaudojimas yra ribojamas. Pirmiausia, panaudojant tok operatori , turi b ti sitikinta, kad panaudojus pakankam kiek interpoliacijos mazg b t gautas geras interpoliuojamos funkcijos artinys. Ta iau, kaip parodo eil paprast pavyzdži , tai dažniausiai garantuoti negalima. Toliau bus pateikti keli iš ši pavyzdži .

S. H. Bernšteinas (1916 m.) nustat , jog interpoliacini Lagrandžo daugianari eiliškumas  $\{L_N(x)\}$ , sudarytas netolydžiai funkcijai  $f(x) = |x|$  atkarpoje  $[-1, 1]$  su vienodai nutolusiais mazgais<sup>9</sup> ( $x_0 = -1, x_N = 1$ ), augant  $N$  neart ja prie funkcijos  $f(x)$ . Dar vienas domus pavyzdys susideda iš to, kad nurodytas interpoliacinis procesas n ra tolydus, ir turi tr k taške  $[-1, 1]$  nors yra turima vien kart tolydžiai diferencijuojama funkcija  $f(x) = (1 + 25x^2)^{-1}$ . Abiejuose išvardintuose atvejuose galioja ribos formul :  $\lim \max |f(x) - L_N(x)| = \infty$ .

Kartais šiuos sunkumus pavyksta veikti specialii interpoliacijos mazg parinkimu arba pereinant prie tam tikr apibendrint daugianari . Ta iau tokiu atveju yra apsunkinami skai iavimai ir augant laipsniui greitai kaupiamos apvalinimo klaidos. D l išvardint priežas i praktikoje tam, kad pakankamai gerai priartinti funkcij , vietoje aukštos eil s interpoliacinio daugianario yra naudojama interpoliacija netolydžiais daugianariais.

Vienas iš toki pavyzdži yra netolydi tiesin interpoliacija. Bendru atveju atkarpa  $[a, b]$  su taškais  $a = \langle_0 < \langle_1 < \dots < \langle_n = b$  yra padalinama tam tikr skai i dali , kuriuose ir sudaromi nauji

<sup>9</sup> Mazgas yra tolydžiai deformuoto apskritimo atvaizdavimas trimat je Euklidin je erdv je -  $\mathbb{R}^3$

interpoliacijos daugianariai. Tokiu būdu gauti daugianariai (dažniausiai to paties laipsnio) duoda funkcijos  $f(x)$  interpoliaciją visoje  $[a, b]$  atkarpoje, kuri bendru atveju negarantuoja tolydaus perėjimo nuo vienos dalies prie kitos ir gali būti netolydi, jeigu taškai  $x_i, i=1, \dots, n-1$  nėra traukiami interpoliacijos mazgais. Tokio tipo interpoliacija naudojama, jeigu reikalaujama atstatyti funkciją su tam tikru tolydumo laipsniu. Bendru atveju, vairios lentelės yra sudaromos su tokiu žingsniu, kad tarpinės funkcijos reikšmės (su tam tikru tikslumu) galima būtų išreikšti tiesine arba kvadratine interpoliacija. Tam, kad tolydžiai atstatyti lentelę užduot funkciją reikia padidinti jos sudarančio daugianario laipsnį, o likusius laisvuosius kintamuosius nustatyti remiantis sąlyga, kad daugianaris susijungimas yra tolydus. Tokiu būdu gaunamos netolydžios daugianarės (polinominės) funkcijos su vientisa struktūra (sudaryta iš daugianarių turinčių tą patį laipsnį), kurios vadinamos splaino funkcijomis arba tiesiog slainais. Pats paprasčiausias ir istoriškai pats seniausias splaino pavyzdys – laužtas.

Jeigu paimsime lanką  $l$  liniuot, vieno jo galą užfiksuosime pradiniam taške  $(x_0, f(x_0))$ , o kitas dalis pritvirtinsime ant duoto taško, tuomet pagal Bernulio – Eilerio taisyklę išlenktos liniuotės ašies diferencialinė lygtis turi pavidalą:

$$EIS''(x) = -M(x), \quad (2.2)$$

čia  $S''(x)$  – antroji linkio išvestinė;

$M(x)$  - išlinkimo momentas, kuris keičiasi tiesiškai nuo vieno atramos taško prie kito;

$EI$  - kietumas.

Apskaičiavus lygties integralą gaunama funkcija  $S(x)$ , kuri aprašo liniuotės profilį. Tai kubinis daugianaris tarp dviejų kaimyninių atramos taškų ir du kartus tolydžiai diferencijuojama funkcija visoje integruojamoje dalyje. Būtinai šio uždavinio sąlyga yra tai, jog turi būti užduotos kraštinės sąlygos  $S''(x_0) = S''(x_N) = 0$ .

Funkcija  $S(x)$  yra tolydaus splaino pavyzdys. Ji yra priskiriama prie kubinio splaino, kurie pasižymi daugeliu unikalių savybių, garantuojančių priglodinimo tikslumą.

Skirtingai nei Lagrandžo interpoliaciniuose daugianariuose, kubinio interpoliacinio splaino eiga tolydžiam tinkle (sudarytame iš mazgų) priveda prie interpoliuojamos tolydžios funkcijos, o kai yra gerinamos diferencialinės funkcijos savybės tuomet priglodinimo tikslumas auga. Tokiu atveju kubinis splainas funkcijai  $f(x) = (1 + 25x^2)^{-1}$  (iš Rung's pavyzdžio) tinkle, kuriame yra  $N = 6$  mazgų duoda tos pačios eilės paklaidą, kaip ir daugianaris  $L_5(x)$ , tačiau kai  $N = 21$ , ji pasidaro itin maža ir nelabai reikšminga.

Kubini spline sudarymo algoritmai yra gana paprasti ir efektyviai realizuojami kompiuterini technologij pagalba, o apvalinimo klaid taka pasidaro nereikšminga (O'Neill C., 2002).

Be to, kubiniai splineai pasižymi domiomis savyb mis, kurios susijusios su faktoriumi, kad kreiv s profilis, kuris eina per užduotus taškus su kraštin mis s lygomis  $S'(x_0) = f'(x_0)$  ir  $S'(x_N) = f'(x_N)$  arba  $S''(x_0) = S''(x_N) = 0$ , tampa tokios formos, kur potenciali energija yra minimali. Tiesiniame priglodinime tai yra išreiškiama tokiu b du:

$$\int_a^b [S''(x)]^2 dx \leq \int_a^b [f''(x)]^2 dx, \quad (2.3)$$

Lygyb galioja tik tuo atveju, kai  $f(x) = S(x)$ .

Funkcijos interpoliacijos splineais ir optimalios tiesini funkcij aproksimacijos, kur tiksl s sprendiniai yra splineo funkcijos tyrimai s lygojo dviej spline teorijos krypt i atsiradim : algebrin s ir variacin s.

Pirmuoju atveju splineai yra suprantami kaip tolydžios iš daugianario dali sudarytos funkcijos, turin ios vientis strukt r. Prie toki funkcij yra priskiriami  $L -$  splineai, kurie gaunami radus tiesin s diferencialin s lygties  $LS(x) = 0$  sprendin . ia kubini spline atveju  $L = d^4 / dx^4$ , o  $S(x_i) = f(x_i)$ . Aproksimacijos uždavinio sprendimas bei aproksimacini splineo s vybi analizavimas susiveda algebrini sistem tiesini funkcij analizavim .

Antruoju atveju, variacin je teorijoje, splineai yra suvokiami kaip Gilberto erdvi elementai, kurie minimizuoja tam tikrus funkcionalus, o taip pat yra analizuojamos sprendini savyb s.

Spline nagrin jim bei jau min t nauj krypt i atsiradim labiausiai takojo dvi esmin s priežastys:

- 1) geras spline sutapatinimas su aproksimuojamais objektais;
- 2) paprastas spline algoritm sudarymas, naudojant informacines technologijas.

Šios dvi savyb s galioja vienma iams splineams. Dvima iu atveju apsiribojama funkcijomis su  $l$  steline strukt ra. Tai tokios funkcijos, kuri erdv yra padalinama  $l$  steles (vienma iu atveju tai sta iakampis, trikampis, o daugiamia iu – gretasieniai, piramid s). Kiekvienoje  $l$  stel je funkcija tam tikra prasme nustatoma vienodais metodais su tolydumo s lyga ties  $l$  steli ribomis. Kai yra interpoliuojama funkcija turinti daug kintam j , splineams, skirtingai nei daugianariams, neiškyla dideli sunkum su egzistavimo ir vienaties sprendiniu.

Erdv ms, kurios yra padalintos gretasienius, daugiamia i spline pl tros tendencija remiasi vienma i spline pl tinio tyrimu, kas garantuoja atitikamumo s vybi išsaugojim ir algoritmiškum , o daugelyje uždavini ir ekstremalias savybes.

Taigi spline funkcijos (kaip funkcijos priartinimo aparato) panaudojimas skaitinėje analizėje parodo, kad visais žinomais atvejais gali būti pasiekti geri palyginimo, su klasikiniu daugianario aparatu, rezultatai. Vienuose uždaviniuose per jį prieš spline'ą lygoja rezultatų tikslumą, kitais atvejais sumažėja skaitiniai metodai, o trečiu atveju yra pasiekiami abu efektai kartu. Spline'ų pritaikymo teorijoje pasisekė išspręsti tokius uždavinius, kuriuos kitais metodais išspręsti buvo tiesiog ne manoma. Toliau trumpai aptarsime šiuos uždavinius.

Yra žinoma geometrinės informacijos suvokimo ir saugojimo problema varioose srityse: technika, architektūra, kartografija ir pan. Tradiciškai daugiau ar mažiau sudėtingose situacijose šis uždavinys yra sprendžiamas objekto pavaizdavimu būdu arba procesu plokštumoje, grafiko ar brėžinio pavidalu. Kadangi pavaizdavimo mastelis yra ribojamas, šis metodas iš principo negali garantuoti reikiamo tikslumo visais atvejais. Tais atvejais yra naudojami spline'ai. Jie leidžia išsaugoti geometrinę informaciją skaitinėje formoje su bet kokių tikslumu. Kai naudojant tam tikrus programos paketus yra apdorojama informacija, spline'ų panaudojimas leidžia metodologiniu pagrindu išplėtoti matematinę mechaninę grafikos programinę rangą (pvz. graf. sudarymo metodai). Taigi spline'ų nauda yra minimas ir rodomas daugelyje šaltinių, todėl jis yra labai plačiai naudojamas ir šiais laikais. (C... .., ... .., ... ..: 1980), (Sankar V.P., Ferrari L.A., 1998), (O'Neill. C., 2002), (Stanford Exploation Project., 2000)

Ypač plačiai spline'ai naudojami technikoje. Jis skirtas detali paviršiaus matematiniam modeliavimui bei sudėtingos agregat formos, tokios kaip laivo korpusas, lengvųjų automobilių kėbulų sudarymui. Tokie matematiniai modeliai tapo būtini sudarant produktų projektavimo automatizacijos sistemas (remiantis informacinėmis technologijomis) bei technologiniam pasiruošimui gamybai, skaitant programų sudarymą prietaisams su skaitmeniniu programiniu valdymu. Tad spline'ų panaudojimas visuose srityse vis auga.

## 2.2 Spline erdvė

Šiame skirsnyje bus apžvelgtos pagrindinės spline'ų teorijos savybės, kurios reikalingos tiriant spline'ų aproksimaciją. Pateiktos funkcijos priglodinimo uždaviniuose, kurie priskiriami tiesiniams rinkiniams, aproksimuojamoji funkcija dažniausiai yra randama tarp šio rinkinio baigtinės erdvės elementų. Funkcijos tiesiniuose rinkiniuose tokia erdvė yra daugianari erdvė, kurios eilėnraukštesnė nei užduotoji. Toliau bus parodyta, kad užduoto laipsnio bei glodumo spline'ai su fiksuotais mazgais taip pat pasižymi baigtinės erdvės savybe ( ... .., ... .., 1980).

Tegu atkarpoje  $[a, b]$  yra užduoti taškai  $\Delta: a = x_0 < x_1 < \dots < x_N = b$ , kuriuose funkcija yra netolydi. Skaičius  $k \geq 0$ , tenkinantis sąlygą  $C^k = C^k[a, b]$ , bus laikomas daugianariu, kuris yra vieną kartą tolydžiai diferencijuojamas funkcijos atkarpoje  $[a, b]$ , o per  $C^{-1}[a, b]$  dali sudaryt tolydži funkcijų daugianar su pirmosios eilės trūkio taškais.

**Apibr žimas.** Funkcija  $S_{n, \hat{\epsilon}}(x)$  yra vadinama  $n$  laipsnio bei defekto  $\hat{\epsilon}$  splainu ( $\hat{\epsilon}$  - sveikasis skaičius,  $0 \leq \hat{\epsilon} \leq n+1$ ) su mazgais tinkle  $\Delta$ , jeigu

- a) Kiekvienoje atkarpoje  $[x_i, x_{i+1}]$  funkcija  $S_{n, \hat{\epsilon}}(x)$  yra  $n$  laipsnio daugianaris. Kuris užduodamas tokiu pavidalu:

$$S_{n, \hat{\epsilon}}(x) = \sum_r^n a_r^i (x - x_i)^r, \text{ kai } x \in [x_i, x_{i+1}], i = 0, \dots, N-1 \quad (2.4)$$

- b)  $S_{n, \hat{\epsilon}}(x) \in C^{n-\hat{\epsilon}}[a, b]$ .

Splaino apibr žimas galioja ir visoje ašyje, jeigu  $a = -\infty, b = +\infty$ . Kiekvienoje atkarpoje  $[x_i, x_{i+1}]$  splainui, galioja ne tik formulė (2.4), bet ir :

$$S_{n, \hat{\epsilon}}(x) = \sum_r^n b_r^i (x - x_{i+1})^r, i = 0, \dots, N-1. \quad (2.5)$$

Ši formulė (2.5) yra naudojama pusašyje  $(-\infty, x_1]$ , o pusašyje  $[x_{N-1}, \infty)$  naudojama tik formulė (2.4).

Taigi, splainas  $S_{n, \hat{\epsilon}}(x)$  turi tolydžias  $n - \hat{\epsilon}$  eilės išvestines. Splaino išvestinys, kuri eilė yra aukštesnė, turi trūkios taškuose  $x_i, i = 1, \dots, N-1$ . Kad funkcija būtų apibr žta, reikia užduoti atitinkamą sąlygą, kad  $S_{n, \hat{\epsilon}}^{(r)}(x), r > n - \hat{\epsilon}$  nra trūkios iš dešinės.

$$S_{n, \hat{\epsilon}}^{(r)}(x_i) = S_{n, \hat{\epsilon}}^{(r)}(x_i + 0), r = n - \hat{\epsilon} + 1, \dots, n; i = 1, \dots, N-1. \quad (2.6)$$

Splainus, kurie tenkina šią apibr žimą žymime  $S_{n, \hat{\epsilon}}(\Delta)$ . Yra aišku, kad šiam daugianariui priskiriami ir  $n$  laipsnio splainai su defektu  $\hat{\epsilon}_1 < \hat{\epsilon}$ , jeigu  $n_1 - \hat{\epsilon}_1 \geq n - \hat{\epsilon}$  ir daugianariai yra ne aukštesnės eilės negu  $n$ . Kadangi atliekant paprastą elementų iš  $S_{n, \hat{\epsilon}}(\Delta)$  sumavimo operacijų arba dauginant juos iš tikrinių skaičių yra tenkinamos, aukščiau minėtų sąlygų, todėl tai vadinama tiesiniu daugianariu arba tiesine erdve.

Paprastiausias splaino pavyzdys yra vienmatė Hevisaido (dar vadinama jungimo funkcija ir yra naudojama signalų teorijoje) funkcija:

$$\Theta(x) = \begin{cases} x^n, & \text{kai } x \geq 0, \\ 0, & \text{kai } x < 0. \end{cases} \quad (2.7)$$

Su šia funkcija yra susijusi laiptuota funkcija

$$x_+^n = x^n \Theta(x) = \begin{cases} x^n, & \text{kai } x \geq 0, \\ 0, & \text{kai } x < 0. \end{cases} \quad (2.8)$$

Funkcijos  $\Theta(x)$  ir  $x_+^n$  yra splainai, atitinkamai nulinio ir  $n$  laipsnio su defektu 1 ir vieninteliu mazgu nuliniame taške. Tam, kad aprašyti splainus reikia nors trumpai aptarti su jais susijusias laiptuotas funkcijas  $(x-x_i)_+^{r'}$ , kurios taip pat susijusios su taškais tinkle  $\Delta$ . Kai  $n-\hat{+}1 \leq r' \leq n$  jos priklauso daugianariui  $S_{n,\hat{+}}(\Delta)$

### 2.3 Pirmos eilės splainas

Išanalizavus ir išsiaiškinus kas yra splain erdvė galima išanalizuoti pagrindines splain savybes. Pats paprasčiausias splaino variantas tai pirmos eilės splainas (Cox, 1970; .., 1980). Tai splaino tipas, kurio paprastume slypi specifinis splain savybės.

$$S_1(x) = f_i \frac{x_{i+1}-x}{h_i} + f_{i+1} \frac{x-x_i}{h_i}, \quad (2.9)$$

$$S_1(x) = f_i + \frac{x-x_i}{h_i} (f_{i+1} - f_i). \quad (2.10)$$

Kartais, norint išskirti interpoliuojamą funkciją vietoj  $S_1(x)$  yra žymima  $S_1(f, x)$ . Tam, kad splainas būtų išsaugotas reikia turėti  $2(N+1)$  taškų. Skaičiavimus patogiau atlikti tokia seka: pirma randamas  $u_i = (f_{i+1} - f_i) / h_i$ , vėliau  $S_1(x) = f_i + (x-x_i)u_i$ . Tuomet, kad išskaičiuoti splain viename taške reikia atlikti šešias aritmetines operacijas. Jeigu reikia rasti reikšmes tos pačios atkarpos  $[x_i, x_{i+1}]$  kituose taškuose, tuomet reikia papildomai atlikti tik po tris operacijas kiekvienam taškui.

Jeigu  $f_i$  reikšmės yra užduotos su paklaidomis, kurios yra mažesnės už  $\nu$ , tuomet reikšmės  $S_1(x)$  yra randamos su tokiu pačiu tikslumu, o išvestinės  $S_1'(x)$  reikšmės yra randamos su tikslumu  $2\nu / h_i$ . Sudingtesnis ir netiesinis yra Hermito splainas.

### 2.4 Hermito splainas

Hermito splainas yra kubinis, todėl tarkime, kad tinklo mazguose  $\Delta: a = x_0 < x_1 < \dots < x_N = b$  yra užduotos tam tikros funkcijos  $f(x)$  ir jos išvestinės  $f'(x)$  reikšmės:  $f_i = f(x_i), f'_i = f'(x_i), i = 0, 1, \dots, N$ .

Kubiniu interpoliaciniu splainu su defektu 2 (hermito kubiniu splainu) vadinsime funkciją  $S_{3,2}(f; x) \equiv S_{3,2}(x)$ , kuri tenkina šlygas:

1) Kiekviename intervale  $[x_i, x_{i+1}]$  turi galioti:

$$S_{3,2}(x) = a_{i0} + a_{i1}(x - x_i) + a_{i2}(x - x_i)^2 + a_{i3}(x - x_i)^3;$$

$$2) \quad S_{3,2}(x_i) = f_i, \quad S'_{3,2}(x_i) = f'_i, \quad i = 0, 1, \dots, N.$$

Akivaizdu, kad  $S_{3,2}(x) \in C^1[a, b]$ . Antroji kubinio Ermito splaino išvestin turi tr kius tinklo mazguose  $\Delta$ . Remiantis interpoliacijos s lygomis, tam kad rasti koeficientus  $a_{ir}, r = 0, 1, 2, 3$  su skirtingais  $i$  yra naudojama lyg i sistema:

$$\begin{cases} S_3(x_i) = f_i \\ S_3(x_{i+1}) = f_{i+1} \\ S'_3(x_i) = f'_i \\ S'_3(x_{i+1}) = f'_{i+1} \end{cases} \quad (2.11)$$

Išspendus ši sistem gauname:

$$S_3(x) = \{\}_1(t)f_i + \{\}_2(t)f_{i+1} + \{\}_3(t)f_i h_i + \{\}_4(t)f'_{i+1} h_i, \quad (2.12)$$

kur

$$\begin{aligned} \{\}_1(t) &= (1-t^2)(1+2t), & \{\}_2(t) &= t^2(3-2t), & \{\}_3(t) &= t(1-t)^2, \\ \{\}_4(t) &= -t^2(1-t), & h_i &= x_{i+1} - x_i, & t &= (x - x_i) / h_i. \end{aligned} \quad (2.13)$$

Formul (2.12) yra patogi teoriniams tyrin jimams. Praktiniams splaino radimo taške  $x \in [x_i, x_{i+1}]$  tyrimams, atsižvelgiant atliekam operacij kiek, patogiau naudoti sekan ias formules:

$$S_3(x) = \frac{(x_{i+1} - x)^2 \cdot (2(x_i - x) + h) \cdot y_i}{h^3} + \frac{(x_{i+1} - x)^2 \cdot (x - x_i) \cdot m_i}{h^2} + \frac{(x_{i+1} - x)^2 \cdot (2(x_i - x) + h) \cdot y_{i+1}}{h^3} + \frac{(x_{i+1} - x)^2 \cdot (x - x_i) \cdot m_{i+1}}{h^2}, \quad (2.14)$$

$$m_i = \frac{(y_{i+1} - y_{i-1})}{2h},$$

$$m_i = \frac{(3y_0 + 4y_1 - y_2)}{2h},$$

$$m_N = \frac{(3y_N - 4y_{N-1} - y_{N-2})}{2h}$$

Tam, kad rasti splain viename taške pakanka atlikti 16 aritmetini operacij , iš j 6 „ilgas“. Nesunku pasteb ti, kad splaino radimo algoritmas yra pastovus vedam duomen paklaid atžvilgiu. Ta iau kai yra randamos Ermito splaino išvestin s  $S_3^{(r)}(x), r = 1, 2, 3$  d l duomen paklaid negalima naudoti per daug tankius tinklus. Optimalus tinklo žingsnio dydis splaino išvestin s radimui gali b ti nustatytas po to kai yra vertintos interpoliacijos paklaidos tokiu pat b du kaip ir pirmos eil s splaino atveju. Taigi šio kubinio splaino interpoliacija remiasi iš dali

sudarytais daugianariais ir  $j$  analize. Tai yra patogus būdas ne tik sprendžiant teorinius uždavinius, tačiau ir naudojant skaitinius metodus. Tačiau daugeliu atveju efektyvesnis metodas yra kubini spline panaudojimas remiantis B – splineais.

## 2.5 B – splineas

B – splineai yra taip pat kubiniai ir laikomi labai efektyviais. Ypač dažnai jie naudojami tuo atveju, kai interpoliacija reikalauja minimalios kompiuterio atminties. Šiame skirsnyje apžvelgsime vien iš ši spline sudarymo būdų bet galimos ir kitos interpretacijos (Sankar V.P., Ferrary L.A., 1998), (Stanford Exploration Project., 2000).

Papildykime tinkl  $\Delta : a = x_0 < x_1 < \dots < x_N = b$  mazgais  $x_{-3} < x_{-2} < x_{-1} < a, b < x_{N+1} < x_{N+2} < x_{N+3}$ . Šie mazgai beveik nepriklausomi, tik periodiniu atveju turi būti išpildyta sąlyga  $h_{N+i} = h_i$ .

Dabar bus galima sudaryti kubinį B – spline intervalė  $(x_{i-2}, x_{i+2})$ . Jeigu B – splineas yra nelyginio laipsnio, tuomet numeracija atliekama pagal vidurinį mazgą. Ieškomas B – splineas bus žymimas  $B_i(x)$ , bei  $y_p = B_i(x_p)$ ,  $M_p = B_i''(x_p)$ .

Kaip ir bet kokiam kubiniam splineui, visiems  $B_i(x)$  galioja lygtis:

$$M_{p-1} + 2M_p + M_{p+1} = \frac{6}{h_{p-1} + h_p} \left( \frac{y_{p+1} - y_p}{h_p} - \frac{y_p - y_{p-1}}{h_{p-1}} \right) \quad (2.15)$$

kur  $p = i-1, i, i+1$ . Kadangi  $B_i(x) \equiv 0$ , kai  $x \in (x_{i-2}, x_{i+2})$ , tai  $B_i^{(r)}(x_{i-2}) = B_i^{(r)}(x_{i+2}) = 0$ ,  $r = 0, 1, 2$ . Šias sąlygas galima pateikti tokiu būdu:

$$\begin{aligned} y_{i-2} = y_{i+2} = 0, \quad M_{i-2} = M_{i+2} = 0, \\ y_{i-1} = \frac{1}{6} h_{i-2}^2 M_{i-1}, \quad y_{i+1} = \frac{1}{6} h_{i+1}^2 M_{i+1}. \end{aligned} \quad (2.16)$$

Rasti parametrai eliminuojami iš lygties (2.15), tuomet gaunama:

$$\begin{cases} (h_{i-2} + h_{i-1})(h_{i-2} + 2h_{i-1})M_{i-1} + h_{i-1}^2 M_i = 6y_i, \\ (h_{i-2} + h_{i-1})M_{i-1} + (h_{i-1} + h_i)M_i + (h_i + h_{i+1})M_{i+1} = 0, \\ h_i^2 M_i + (h_i + h_{i+1})(2h_i + h_{i+1})M_{i+1} = 6y_{i+1}. \end{cases} \quad (2.17)$$

Tokiu būdu gaunama trijų lygčių sistema keturi parametrai  $(y_i, M_{i-1}, M_i, M_{i+1})$  radimui. Vien iš parametrai galima naudoti savo nuožiūra. Tarkime

$$y_i = \frac{h_{i-1}(h_{i-2} + h_{i-1})(2h_i + h_{i+1}) + h_i(h_i + h_{i+1})(h_{i-2} + 2h_{i-1})}{(h_{i-1} + h_i)(h_{i-2} + h_{i-1} + h_i)(h_{i-1} + h_i + h_{i+1})} \quad (2.18)$$

Tuomet iš lygčių sistemos (2.17) yra randami:

$$\begin{aligned}
M_{i-1} &= 6(h_{i-2} + h_{i-1})^{-1}(h_{i-2} + h_{i-1} + h_i)^{-1}, \\
M_i &= -6(h_{i-1} + h_i)^{-1}[(h_{i-2} + h_{i-1} + h_i)^{-1} + (h_{i-1} + h_i + h_{i+1})^{-1}], \\
M_{i+1} &= 6(h_i + h_{i+1})^{-1}(h_{i-1} + h_i + h_{i+1})^{-1}.
\end{aligned}
\tag{2.19}$$

Formulės (2.16), (2.18), (2.19) pilnai apibūdina  $B_i(x)$  splain atkarpoje  $[x_{i-2}, x_{i+2}]$ . Lentelėje 1 pateiktos šio splaino reikšmės mazguose  $x_p, p = i - 2, \dots, i + 2$ .

**2.1 lentelė . Kubinio B - splaino reikšmės pateiktuose taškuose**

$x$	$B_i(x)$	$B_i(x)$	$B_i(x)$
$x_{i-2}$	0	0	0
$x_{i-1}$	$\frac{h_{i-2}}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$	$\frac{3h_{i-2}}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$	$\frac{6}{(x_i - x_{i-2})(x_{i+1} - x_{i-2})}$
$x_i$	$1 - \frac{1}{(x_{i+1} - x_{i-1})} \left( \frac{h_i^2}{x_{i+1} - x_{i-2}} + \frac{h_{i-1}^2}{x_{i+2} - x_{i-1}} \right)$	$\frac{3}{(x_{i+1} - x_{i-1})} \left( \frac{h_i}{x_{i+1} - x_{i-2}} + \frac{h_{i-1}}{x_{i+2} - x_{i-1}} \right)$	$-\frac{6}{(x_{i+1} - x_{i-1})} \left( \frac{1}{x_{i+1} - x_{i-2}} + \frac{1}{x_{i+2} - x_{i-1}} \right)$
$x_{i+1}$	$\frac{h_{i+1}^2}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$	$-\frac{3h_{i+1}}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$	$\frac{6}{(x_{i+2} - x_i)(x_{i+2} - x_{i-1})}$
$x_{i+2}$	0	0	0

Nesunku parodyti, kad parametro  $y_i$  pasirinkimas, remiantis formule (2.18) užtikrina B – splaino normalumą. Iš tikrųjų, du kubiniai B – splainai su vienu ir tuo pačiu  $(x_{i-2}, x_{i+2})$  gali skirtis tik pastoviais dauginamaisiais. Tam kad jie sutaptų, pakanka sąlygos, kad jie gautų tas pačias reikšmes viename intervalo  $(x_{i-2}, x_{i+2})$  taške. Šis metodas yra taip pat naudojamas kalvų, slėnių dugno paviršiams modeliuoti. Tačiau kaip yra žinoma dažnai sąryšis tarp priklausomų ir nepriklausomų kintamųjų yra nežinomas, o nepriklausomų kintamųjų gali būti ne vienas, tokiu atveju yra naudojami MAR splainai.

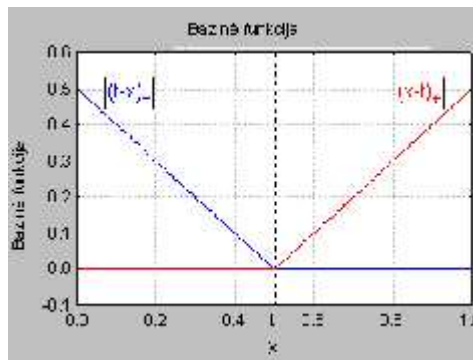
## 2.6 Daugialypiai regresiniai splainai

MAR (angl. *multivariate regression splines*) – splainai tai neparimetrinis procedūra, kurios metu nenaudojamos jokios prielaidos apie funkcionalinius sąryšius tarp priklausomų ir nepriklausomų kintamųjų. Naudojant MAR splainus, remiantis pradiniais duomenimis, yra nustatomos koeficientų ir bazinių funkcijų parinkimo priklausomybės. Erdvė yra dalinama dalis,

turinias regresijos ir klasifikacijos lygtis. Šis faktas MAR – splain panaudojimą daro ypač efektyviu uždaviniams, kuriuose yra dirbama daugiau nei su dvejais nepriklausomais kintamaisiais. Taip pat MAR – splain metodas gali būti labai plačiai panaudojamas duomenų gavimo srityje todėl, kad jis nesiremia prielaidomis apie uždavinio tipą ir nepateikia jokių apribojimų priklausomybei klasei (pavyzdžiui tiesinei ar logistinei) tarp priklausomo ir nepriklausomo kintamojo. Tokiu būdu šis metodas leidžia gauti ypač tikslius modelius net ir tokiais atvejais, kai ryšiai tarp priklausomo ir nepriklausomo kintamojo yra nemonotoniško charakterio ir sudėtinga panaudoti parametrinius modelius.

Kaip jau buvo minėta regresijos uždaviniai yra sudaryti ryšių tarp priklausomo ir nepriklausomo kintamojo nustatymui. Vienas iš regresijos uždavinio pavyzdžių gali būti pateiktas tiriant auto vykius. Paimkime auto vykius, kurie priklauso nuo dviejų aspektų: vairavimas išgirus arba blogos oro sąlygos. Tuomet turime:  $e = r + 0,5f + 2g$ . Čia  $e$  – auto vykių skaičius,  $r$  – tam tikra konstanta,  $f$  – blogos oro sąlygos,  $g$  – vairavimas išgirus. Auto vykių skaičius yra priklausomas kintamasis, kurio reikšmė priklauso nuo  $f$  ir  $g$  (blogos oro, vairavimo išgirus) kintamųjų reikšmių. Pastebime, kad nepriklausomi kintamieji yra imami su tam tikrais dauginamaisiais 0,5 ir 2 atitinkamai. Šie kintamieji yra vadinami regresiniais koeficientais. Kuo didesnis regresinio kintamojo modulis, tuo labiau nepriklausomas kintamasis tampa priklausomas. Tarkime, kad šiame paprastame, išgalvotame pavyzdyje kintamųjų reikšmės buvo normuotos ir privestos prie reikšmių su vidurkiu 0 ir standartiniu nuokrypiu 1, tuomet galima daryti išvadą, kad avarinės situacijos dėl neblaivių vairuotojų kaltės yra 4 kartus tikimesnės. Tuo atveju, jei reikšmės  $n$  yra normuotos, išmatuotos naudojant skirtingus matavimo matavimus bei turi skirtingą mastelį, tuomet regresiniai kintamųjų palyginimas yra nereikšmingas. Šis pavyzdys apie auto vykių skaičius yra tipinis tiesinis regresijos modelio panaudojimo situacija, kurioje yra daroma prielaida, kad egzistuoja tiesinis priklausomybės tarp priklausomo ir nepriklausomo kintamojo. Tiesinė regresija yra priskiriama taip vadinamam parametriniam metodų kategorijai. Šie metodai grindžiami prielaida, kad ryšys tarp priklausomo ir nepriklausomo kintamojo yra apriorinis. Naudojantis neparametriniais metodais yra priešingai. Šiuo atveju prielaida apie tam tikrus ryšius tarp priklausomo ir nepriklausomo kintamojo yra atmetama, tuo tarpu modelio funkcija yra pritaikoma remiantis tik turimais duomenimis. Tad tikslingiausia yra naudoti MAR – splainus, kadangi tai neparametrinis procedūra, kuri nenaudoja jokių tipinių funkcinių ryšių tarp kintamojo. MAR – splainai konstruoja ryšius pagal koeficientų surinkimą ir bazines funkcijas, kurios pilnai nustatomos pagal duomenis. MAR – splaino mechanizmo algoritmas galima sivaizduoti kaip tiesinę regresiją, turinčią trikus. Kontroliniai taškai (vertinami pagal duomenis) nustato konkrečius (labai paprastos) tiesinio lygties veikimo erdvę. Taip pat ieškant sprendimų būtina žinoti, kad MAR

splainai kaip bazines funkcijas naudoja dvipuses funkcijas (2.1 pav.), kurios priartina ryšius tarp priklausomo ir nepriklausomo kintamojo (prediktoriaus).



**2.1 pav. Dviej bazini , dvipusi funkcij grafinis vaizdas**

Paveiksle (2.1) pateiktas dviej paprast bazini funkcij  $(t-x)_+$  ir  $(x-t)_+$  (Hastie, M., 2002, pav 9.9) pavyzdys, kur  $t$  – bazini s funkcijos mazgo taškas (parametro reikšm nustato tr kios tiesin s regresijos nariai). Mazgai (parametro  $t$  reikšm s) yra nustatomi pagal duomenis. Simbolis „+“ kuris seka po išraišk  $(t-x)$  bei  $(x-t)$  reiškia, kad tiriamos tik teigiamos atitinkamos lygties reikšm s. Jei mes turime neigiamas funkcijas reikšmes, tuomet j reikšm yra prilyginama nuliui. Ši situacija yra taip pat pademonstruota aukš iau pateiktame paveiksle. Apžvelgus MAR veikimo mechanizmo algoritm ir pagrindines funkcijos b tina pateikti bei aptarti šio tipo splaino matematin model .

Bazini s funkcijos kombinacijoje su modelio parametrais (kurie vertinami mažiausi kvadrat metodu) yra panaudojamos pradini duomen prognozavimui. Bendra MAR lygtis yra užrašoma tokiu pavidalu (Hastie., M. 2002) :

$$y = f(X) = \beta_0 + \sum_{m=1}^M \beta_m h_m(X) \quad (2.20)$$

Formul je (2.20) y yra nusp jamas funkcijos nepriklausom kintam j reikšm mis arba j kombinacija. Ši funkcija (2.20) yra užduodama parametru  $\beta_0$  ir svertine kel tos bazini funkcij  $h_m(X)$  suma. Kitaip tariant pagrindinis šio modelio veikimo principas yra grindžiamas reikiamos svertin s (bazini funkcij ) sumos parinkimu iš bendro bazini funkcij rinkinio, dengian io visas kiekvieno nepriklausomo kintamojo reikšmes (t.y rinkinys bus sudarytas iš vienos bazini s funkcijos ir parametro  $t$  kiekvieno kintamojo kiekvienai reikšmei). MAR – splain algoritmas randa erdv je visus nepriklausomus ir priklausomus kintamuosius bei s ryšius tarp j . Paieškos procese modelio papildymas bazini mis funkcijomis iš bendro rinkinio auga iki tol, kol nebus maksimizuotas bendras mažiausi kvadrat parinkimo kokyb s kriterijus. Atliekant šias operacijas algoritmas atrenka pa ius reikalingiausius nepriklausomus kintamuosius bei taip pat atskleidžia reikšmingiausius ryšius tarp j .

Visi išvardinti interpoliacijos metodai gali būti panaudojami ne tik kalv bei slėni paviršiams modeliuoti, bet ir klimato, užterštumo bei geologini žemėlapi sudarymui. Tačiau ne tik vairių spline funkcijų pagalba, bet ir kriging interpoliacijos metodais galima gauti norimus rezultatus (ieškomas pavirši reikšmes).

## 2.7 Ordinarusis krigingas

Ordinarusis krigingas pasižymi tuo, kad jis yra tiesinis ir vidurkis yra nežinomas. Šio tipo krigingo svoriai  $\{w_i\}$  priklauso nuo išmatuotiems taškams pritaikyto modelio, atstumo iki prognozės vietos ir erdviniai ryšiai tarp aplink prognozės vietą išmatuotų reikšmių. Prognozės skaičiuojama kaip tiesinis žinomų duomenų reikšmių kombinacija (svertinis vidurkis):

$$\hat{Z}(s) = \sum_{i=1}^n w_i Z(s_i) \quad (2.21)$$

Yra laikoma, kad žinomuose taškuose apskaičiuotos reikšmės yra tikslios ir neturi paklaidų. Kai turime ordinarųjį krigingą, tuomet artimesiems taškams yra priskiriami didesni svoriai, o paklaidos verčiai priklauso tik nuo matuojamų taškų išsidėstymo ir nepriklauso nuo duomenų reikšmių.

Kaip ir taikant bet kokį krigingo metodą yra naudojamos tam tikros sąlygos bei prielaidos (Geostatistics, 2003).

Prielaidos taikant ordinarųjį krigingą:

- 1) Bdingas regioninio kintamojo stacionarumas
- 2) Stebėjimų skaičius yra pakankamas tam, kad būtų vertinta variograma.

Taip pat yra pateiktos sąlygos, kada gali būti pritaikomas ordinarusis krigingas.

- 1) Vidurkis  $E[Z(s)] = \mu$  yra nežinomas, tačiau pastovus.
- 2) Funkcijos  $Z(s)$  semivariograma yra žinoma ir aprašoma formule:

$$\gamma(s_1, s_2) = E\left\{\left(Z(s_1) - Z(s_2)\right)^2\right\} \quad (2.22)$$

Nustatant taškų reikšmes naudojantis ordinariuoju krigingu, gali atsitikti taip, kad kai kurios reikšmės bus didesnės, o kai kurios mažesnės už tikrąsias. Tačiau iš pirmos eilės stacionarumo sąlygos žinoma, jog skirtum tarp numatytų ir tikrųjų reikšmių vidurkis turi būti nulis. Ši sąlyga yra laikoma pagrindiniu ordinaraus krigingo apribojimu. Šis apribojimas galima taip pat pateikti kaip reikalavimą, kad visiems išmatuotiems taškams priskirti svoriai  $w_i$  suma būtų lygi vienetui  $\sum_{i=1}^n w_i = 1$ .

Taip pat taikant ordinarj kriging b tina žinoti ir antros eil s stacionarumo s lyg , pagal kuri dispersija kiekviename taške yra baigtin ir vienoda visose visoje srityje, o kovariacijos strukt ra priklauso tik nuo atstumo tarp poros tašk .

Sistemin s paklaidos neb vimo s lyga yra naudojama kartu su „optimalaus“ interpoliavimo prielaida (prognoz s dispersija yra kiek manomai maža) ir yra išreiškiama taip:

$$\begin{bmatrix} x(\mathbf{s}_1, \mathbf{s}_1) & \cdots & x(\mathbf{s}_1, \mathbf{s}_n) & 1 & f_j(\mathbf{s}_1) & \cdots & f_j(\mathbf{s}_1) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x(\mathbf{s}_n, \mathbf{s}_1) & \cdots & x(\mathbf{s}_n, \mathbf{s}_n) & 1 & f_j(\mathbf{s}_n) & \cdots & f_j(\mathbf{s}_n) \\ 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\ f(\mathbf{s}_1) & \cdots & f(\mathbf{s}_n) & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ f_p(\mathbf{s}_1) & \cdots & f_p(\mathbf{s}_n) & 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} \} _i \\ \cdots \\ \} _n \\ \Psi_0 \\ \Psi_1 \\ \cdots \\ \Psi_p \end{bmatrix} = \begin{bmatrix} x(\mathbf{s}_1, \mathbf{s}_0) \\ \cdots \\ x(\mathbf{s}_n, \mathbf{s}_0) \\ 1 \\ f(\mathbf{s}_0) \\ \cdots \\ f_p(\mathbf{s}_0) \end{bmatrix} \quad (2.23).$$

Šio krigingo pagrindas yra statistiniai modeliai, kuriuose yra atsižvelgiama erdvin autokoreliacij (statistin ryš tarp atramos tašk ). Šia savybe taip pat pasižymi ir universalus krigingas, be to abu šie metodai gali b ti panaudojami interpoliuojam reikšmi erdvi sudaryme, o taip pat atliktos interpoliacijos patikimumo bei tikslumo nustatymui. Dažniau yra naudojamas, katik aprašytas, ordinarusis krigingo metodas, ta iau jei turime duomenis, kuriuose galima pasteb ti tam tikr reikšmi dominavimo tendencij tuomet naudojamas universalus krigingas.

## 2.8 Paprastasis krigingas

Prieš pradedant nagrin ti kelias populiariausias krigingo r šis bus pateiktos kelios bendros formul s apiman ios visas krigingo r šis. Bendra krigingo interpoliatori formul - svertin duomen suma (Geostatistics., 2003).

$$\widehat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \} _i Z(\mathbf{s}_i) \quad (2.24)$$

ia  $Z(\mathbf{s}_i)$  yra i-tojoje vietoje išmatuota reikšm ;

$\mathbf{s}$  – nagrin jamos srities taškai;

$\} _i$  - nežinomas i- tojoje vietoje išmatuotas reikšm s svoris;

$\widehat{Z}(\mathbf{s}_0)$  - prognoz pradiniame taške  $\mathbf{s}_0$  ;

o  $n$  – išmatuot reikšmi skai ius.

Taip pat svarbu pamin ti bendr j atsitiktinio lauko model  $Z(\mathbf{s}) = \sim(\mathbf{s}) + v(\mathbf{s})$  (šis modelis išsamiai išnagrin tas 1.1 skirsnyje).

Kaip jau buvo min ta pirmoje dalyje, paprastasis kringas yra tiesinis su žinomu trendu. Taip pat tai yra matematiškai paprasčiausia kringo r šis. ia daroma prielaida, kad atsitiktinio lauko tikimyb turi b ti žinoma ir remiasi kovariacine funkcija. Ta iau daugelyje atveju yra pateikiami tokie uždaviniai, kur nei atsitiktinio lauko tikimyb nei kovariacij n ra žinomi iš anksto. Paprastasis kringas skiriasi nuo kit r ši tuo, kad skai iavimams n ra naudojamas Lagrandžo daugiklis. Be to paprastam kringui negalioja apribojimas, kad svori suma turi b ti 1, tai yra  $\sum_{i=1}^n \lambda_i \neq 1$ . Ta iau skai iuojant taško reikšm , nuokrypius nuo ši svori reikia kompensuoti atsižvelgiant žinom vidurk .

Nesud tinga parodyti, kad BLUP (geriausia tiesin nepaslinkta prognoz ) taške  $\mathbf{s}_0 \in D$  paprastojo kringo atveju yra (Duninskas K., 2003):

$$\begin{aligned} \mathbf{a}^T &= \mathbf{C}_0^T \mathbf{V}^{-1}, \\ a_0 &= \tilde{z}(\mathbf{s}_0) - \mathbf{C}_0^T \mathbf{V}^{-1} \boldsymbol{\mu}_n \end{aligned} \quad (2.25)$$

kur

$$\begin{aligned} \mathbf{C}_0^T &= (C(\mathbf{s}_0 - \mathbf{s}_1), C(\mathbf{s}_0 - \mathbf{s}_2), \dots, C(\mathbf{s}_0 - \mathbf{s}_n)), \\ \boldsymbol{\mu}_n &= (\tilde{z}(\mathbf{s}_1), \tilde{z}(\mathbf{s}_2), \dots, \tilde{z}(\mathbf{s}_n))^T \end{aligned}$$

Tuomet  $MSPE(\hat{Z}(\mathbf{s}_0)) = C(\mathbf{0}) - \mathbf{C}_0^T \mathbf{V}^{-1} \mathbf{C}_0$ . MSPE – vidutin kvadratin paklaida. O prognoz s klaida (dispersija) :  $\sum_{i=1}^n \lambda_i \chi(\mathbf{s}_i - \mathbf{s}_0) + \Psi = \dagger^2(\mathbf{s}_0)$ , kur  $\Psi$  Lagranžo daugiklis, kuris priklauso nuo  $\mathbf{s}_0$ .

Taip pat paprastojo kringo lyg i sistema gali b ti užrašyta tokiu pavidalu :

$$\begin{bmatrix} \chi(\mathbf{s}_1, \mathbf{s}_1) & \cdots & \chi(\mathbf{s}_1, \mathbf{s}_n) \\ \cdots & \cdots & \cdots \\ \chi(\mathbf{s}_n, \mathbf{s}_1) & \cdots & \chi(\mathbf{s}_n, \mathbf{s}_n) \end{bmatrix} \times \begin{bmatrix} \lambda_1 \\ \cdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} \chi(\mathbf{s}_1, \mathbf{s}_0) \\ \cdots \\ \chi(\mathbf{s}_n, \mathbf{s}_0) \end{bmatrix} \quad (2.26)$$

Paprastojo kringo klaidos dispersija skai iuojama pagal formul :  $\dagger^2(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i \chi(\mathbf{s}_i - \mathbf{s}_0)$ .

Taikant š kringo metod yra naudojamosi trimis praktin mis prielaidomis.

- 1) laukas stacionarus pla i ja prasme.
- 2)  $\tilde{z}(x) = 0$
- 3) Kovariacin funkcija:  $c(x, y) = \text{cov}(Z(x), Z(y))$

Paprasto kringo metodas labai retai yra taikomas praktikoje, kadangi dažniausiai mes nežinome vidurkio funkcijos. Kai vidurkio funkcija yra nežinoma tuomet specialistai dažniausiai naudoja ordinar arba universal kringing .

## 2.9 Universalus kringas

Kaip jau ir buvo min ta anks iau, tiesini stochastini metod atsitiktin komponent (bendras atsitiktinis lauko modelis)  $Z(\mathbf{s})$  yra išreiškiama tokia lygtimi:  $Z(\mathbf{s}) = \tilde{v}(\mathbf{s}) + v(\mathbf{s})$ . Tarkime, kad dabar

$$\tilde{v}(\mathbf{s}) = \sum_{k=1}^N f_k(\mathbf{s}) \quad (2.27)$$

ia funkcijos erdvin s koordinat s yra žinomos, s gali b ti užrašoma  $s = (s_1, s_2)$ ; funkcija  $f_k(s)$  gali b ti išreiškiama naudojantis tiesiniu trendu su atitinkamomis koordinat mis:  $f_1((s_1, s_2)) = 1$ ,  $f_2((s_1, s_2)) = s_1$ ,  $f_3((s_1, s_2)) = s_2$ ; jeigu turime trendo kvadrat , tuomet atsiranda papildomos funkcijos:  $f_4((s_1, s_2)) = s_1^2$ ,  $f_5((s_1, s_2)) = s_2^2$ ,  $f_6((s_1, s_2)) = s_2 s_1$ .

Tarkime, kad mes tiriamo kvie i lauko derlingum , tuomet funkcija  $f_k(s)$  gali aprašyti papildom regionuose esan i informacija:

1. tr š kiek ,
2. dirvožemio matavimus, tokius kaip molis arba humuso kiekis
3. vandens kiek ir t.t.

Manoma, kad  $v(s)$  yra b dingas stacionarumas ir tenkinamos sekan ios s lygos:

$$\text{var}(v(\mathbf{s}_i) - v(\mathbf{s}_j)) = 2\lambda(\mathbf{s}_i - \mathbf{s}_j), \quad (2.28)$$

$$\widehat{Z}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i \tilde{v}(\mathbf{s}_i) = \sum_{i=1}^n \sum_{k=1}^N \lambda_i f_k(\mathbf{s}_i).$$

vertinys yra nepaslinktas, tada ir tik tada, jei

$$\sum_{i=1}^n \lambda_i f_k(\mathbf{s}_i) = f_k(\mathbf{s}_0), \quad \forall k = 1, \dots, N. \quad (2.29)$$

Svoriai yra randami, kaip prasta, minimizuojant prognoz s dispersij . Tai yra daroma sprendžiant tiesin lyg i sistem :

$$\begin{cases} \chi(\mathbf{s}_0, \mathbf{s}_j) = \sum_{i=1}^n \lambda_i \chi(\mathbf{s}_i - \mathbf{s}_j) + m_0 + \sum_{k=1}^N m_k f_k(\mathbf{s}_j), \quad j = 1, \dots, n, \\ f_k(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i f_k(\mathbf{s}_i), \quad k = 1, \dots, N \\ \sum_{i=1}^n \lambda_i = 1. \end{cases} \quad (2.30)$$

ia  $m_0, m_1, \dots, m_N$  yra  $N+1$  Lagrandžo daugikliai;

$\chi(\mathbf{s}_i - \mathbf{s}_j)$  yra semivariagramos reikšm s tarp tašk  $\mathbf{s}_i$  ir  $\mathbf{s}_j$ ,

$\chi(\mathbf{s}_0 - \mathbf{s}_j)$  yra semivariagramos reikšm s tarp tašk  $\mathbf{s}_0$  ir  $\mathbf{s}_j$ .

Universalaus krigingo sistema gali būti pateikiama ir matricine forma:

$$\mathbf{f} = \mathbf{f} \quad (2.31)$$

ia

$$\begin{aligned} &= (\mu_1, \mu_2, \dots, \mu_n, m_0, m_1, \dots, m_N)' \\ &= (X(\mathbf{s}_0 - \mathbf{s}_1), X(\mathbf{s}_0 - \mathbf{s}_2), \dots, X(\mathbf{s}_0 - \mathbf{s}_N), f_1(\mathbf{s}_0), f_2(\mathbf{s}_0), \dots, f_N(\mathbf{s}_0))', \end{aligned}$$

ir

$$\mathbf{f} = \begin{bmatrix} \mathbf{F}' \\ \mathbf{F} \quad \mathbf{0} \end{bmatrix},$$

kai

$$\mathbf{F} = \begin{bmatrix} X(\mathbf{s}_1 - \mathbf{s}_1) & X(\mathbf{s}_1 - \mathbf{s}_2) & \dots & X(\mathbf{s}_1 - \mathbf{s}_n) \\ X(\mathbf{s}_2 - \mathbf{s}_1) & X(\mathbf{s}_2 - \mathbf{s}_2) & \dots & X(\mathbf{s}_2 - \mathbf{s}_n) \\ \vdots & \vdots & \dots & \vdots \\ X(\mathbf{s}_n - \mathbf{s}_1) & X(\mathbf{s}_n - \mathbf{s}_2) & \dots & X(\mathbf{s}_n - \mathbf{s}_n) \\ f_1(\mathbf{s}_1) & f_1(\mathbf{s}_2) & \dots & f_1(\mathbf{s}_n) \\ f_2(\mathbf{s}_1) & f_2(\mathbf{s}_2) & \dots & f_2(\mathbf{s}_n) \\ \vdots & \vdots & \dots & \vdots \\ f_n(\mathbf{s}_1) & f_n(\mathbf{s}_2) & \dots & f_n(\mathbf{s}_n) \end{bmatrix}.$$

ia svoriai  $\mathbf{f}^{-1}$ ;

universalaus krigingo prognozės dispersija  $\hat{\sigma}_{UK}^2$  yra  $\hat{\sigma}_{UK}^2 = \hat{\sigma}^2$ .

Jei vidurkis yra nestacionarus, tuomet gaunamas aukštesnės eilės modelis. Atlikus šį procesą bus nebe manoma vertinti variogramos. Taip atsitinka, kadangi yra nežinomi  $\mu_1, \mu_2, \dots, \mu_N$ . Dėl šios priežasties negalima apskaičiuoti  $\hat{\sigma}^2(s)$ , kurie yra reikalingi empirinės variogramos skaičiavimui. Jeigu variograma žinoma tuomet galima vertinti vektorius  $\mathbf{f} = (\mu_1, \mu_2, \dots, \mu_N)$  parametrus pasinaudojus mažiausi kvadratų vertiniu:  $\hat{\mathbf{S}}_{gls} = (\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{F})^{-1}\mathbf{F}'\mathbf{\Sigma}^{-1}\mathbf{Z}$ . Šioje lygtyje  $\mathbf{F}$  toks pat kaip ir nurodytas aukščiau,  $\mathbf{Z} = (Z(\mathbf{s}_1), Z(\mathbf{s}_2), \dots, Z(\mathbf{s}_n))$ , o kovariacijų matrica  $\mathbf{\Sigma} = \text{var}(\mathbf{Z})$  yra antros eilės ir susijusi su prieš tai pateikta  $\Gamma$ :

$$\mathbf{\Sigma} = \hat{\sigma}^2 \mathbf{I}_n - \mathbf{\Gamma}, \quad (2.32)$$

ia  $\hat{\sigma}^2$  yra dispersija;

$\mathbf{I}_n$  matrica kurios eilės ir stulpelių skaičius lygus  $n$  (Cressie, N. 1993), (Neuman, S.P., Jacobson E.A. 1984).

Universalusis krigingas yra plačiau naudojamas, tačiau prieš jį naudojant reikia žinoti jo trūkumus bei privalumus.

**Tr kumai:**

- 1) Apskai iuota variograma remiantis liekan proces vertinimu paprastai yra neobjektyvi.
- 2) vertintas liekan procesas  $\hat{v}$  labiau neigiamai koreliuotas nei pradinis procesas  $v$ .
- 3) Yra rodyta, kad apskai iuota universaliojo krigingo metodu dispersija yra mažesnė nei turėtų būti.

**Privalumai:**

- 1) Pradžioje variogramos vertinio poslinkis yra mažas. Tokiu būdu priderinant modelį prie empirinės variogramos mažesni svoriai yra priskiriami didesniems lagams (lagas – skirtumas tarp momentų).
- 2) Kadangi augant skirtumams tarp momentų krigingo svoriai sparčiai mažėja, todėl priderinta variograma vertinama tik mažoje tiriamo taško aplinkoje (tiksliai toje vietoje, kurioje priderinimas geras).

**2.10 Medianinio išlyginimo krigingas**

Kaip ir prieš tai nagrinėjtu universalaus krigingo atveju medianinio išlyginimo krigingo atsitiktinės komponentės pavidas yra toks pat:  $Z(\mathbf{s})$  yra išreiškiama tokia lygtimi  $Z(\mathbf{s}) = \tilde{m}(\mathbf{s}) + v(\mathbf{s})$ . Tačiau šiuo atveju trendas  $\tilde{m}(\mathbf{s})$  yra sudarytas iš bendro vidurkio, eilučių bei stulpelių medianų efektų atitinkamai:  $\tilde{m}(\mathbf{s}) = a + r(x) + c(y), \mathbf{s}(x, y)$ . Šie efektai yra vertinami remiantis sekančiomis formulėmis:

$$\begin{aligned} \hat{a} &= \frac{1}{n} \sum_{i=1}^n Z(s_i), \\ \hat{c}(x) &= \frac{1}{q} \sum_{N_k} Z(s_i), k = 1, 2, \dots, q, \\ \hat{r}(y) &= \frac{1}{p} \sum_{M_l} Z(s_i), l = 1, 2, \dots, p, \end{aligned} \quad (2.33)$$

kur  $N_k$  yra taškai esantys  $k$ -tajame gardelės stulpelyje, o  $M_l$  taškai esantys  $l$  – tojoje gardelės eilutėje.

Medianinio išlyginimo metu gaunamos liekanos yra prognozuojamos teigiant, kad juose nėra trendo (interpoliacija ordinariu krigingu). Prognoz atliekama naudojantis ordinaraus krigingo metodu naudojama (2.21) formulė.

Tuomet gautus prognozes rezultatus sudedame su prieš tai apskaičiuotais bendru, stulpelių bei eilučių efektais. Taip gaunamas interpoliuotas paviršius nagrinjamam reikšmių ribose.

## 2.11 Modeli patikros metodai bei analiz

Vienas iš populiariausių bei dažniausiai naudojamų modeli patikrinimo bei palyginimo metodų yra - kryžminis modelio patikra. Jei neturime nepriklausomo duomenų rinkinio, pagal kurį galėtume vertinti modelį, jį galima patikrinti naudojant tuos pačius išmatuotus taškus, iš kurių modelis buvo sudarytas; Taigi, atliekant kryžminį patikrą (angl. *cross-validation*), trendams ir autokoreliacijos modeliams tikrinti naudojami visi išmatuoti duomenys. Turint mažus duomenų imtį, kaip nagrinėjamo atveju, naudojamas metodas, kurio po vieną pašalinamas kiekvienas taškas (angl. *Leave - One - Out*), ir prognozuojama reikšmė tame taške bei remiantis tikrąja reikšme apskaičiuojama prognozės paklaida. Šis procedūra yra atliekama visiems duomenims, o vėliau randama vidutinė kvadratinė paklaida. Pagrindinės charakteristikos kurios nustato modelio tinkamumą bei tikslumą yra šios:

- 1) vidutinė kvadratinė paklaida (VPK) (angl. *ME-Mean Error*), kurios reikšmė turi būti artima nuliui;
- 2) šaknis iš vidutinės kvadratinės prognozės paklaidos (ŠVKPK) (angl. *RMPSE- Root Mean Squared Prediction Error*), kuris turi turėti kuo mažesnę reikšmę;
- 3) bei vidutinio kvadratinio nuokrypio reikšmė (VKNS) (angl. *MSDR- Mean Squared Deviation Ratio*), kuri turi būti artima vienetui (Rossiter, 2004).

$$(VPK) = \frac{\sum_{i=1}^n \left( Z^*(s_i) - Z(s_i) \right)}{n} \quad (2.34)$$

$$(\text{ŠVKPS}) = \sqrt{\frac{\sum_{i=1}^n \left( Z^*(s_i) - Z(s_i) \right)^2}{n}} \quad (2.35)$$

$$(VKNS) = \frac{\sum_{i=1}^n \left[ \left( Z^*(s_i) - Z(s_i) \right) / \dagger(s_i) \right]^2}{n} \quad (2.36)$$

- čia  $Z(s_i)$  - prognozuota reikšmė,  
 $Z(s_i)$  - stebima reikšmė,  
 $\dagger(s_i)$ ,  $s_i$  matavimų vietos standartinė prognozės klaida.

Taip pat modelio tinkamumas bei jo polinominių lygčių koeficientai yra nustatomi remiantis dispersine ANOVA analize. ANOVA tiria priklausomybę tarp nepriklausomų ir priklausomų kintamųjų. Ji parodo, ar modelyje yra su priklausomu kintamuoju susijusi regresorių. Faktorius neturintis tokios priklausomam kintamajam, jei skirstiniai visoms grupoms bus identiški. Tai bus

tuomet, kai sutaps vis grupi vidurkiai, nes dispersijos yra vienodos, o skirstiniai normalieji. Tod norint nustatyti, ar faktorius daro tak atsakui, tikrinama atitinkama nulin hipotez . Jeigu  $p$  reikšm didesn už 0,05 tuomet nulin hipotez apie vidurki lygyb yra priimama, regresijos modelio tinkamumas labai abejotinas. Kadangi prognoz yra atliekama pagal koordinates ji b t netiksli ir nepatikima. Priešingu atveju, jeigu  $p$  reikšm mažesn už 0,05, tai gauname patvirtinim , jog modelis yra tinkamas tolesniam tyrimui ir rezultatus galima laikyti patikimais.

Dispersin s analiz s rezultatus gauname pasinaudoj  $F = \frac{A}{B}$  formule, kur  $A$  vidin kvadrat suma, kuri atspindi duomen kintamum tarp faktoriaus lygi ,  $B$  atspindi kintamum grup s viduje. Šios sumos turi atitinkamai  $(K - 1)$  bei  $(N - K)$  laisv s laipsni . Šios variacijos yra apskai iuojamos pagal sekan ias formules:

$$A = \sum_i n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1),$$

$$B = \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 / (N - K).$$

Tre iasis ne mažiau naudojamas modelio tikrinimo metod yra Akaik s informacijos kriterijus. Akaike informacijos kriterijus (Akaike 1973) detaliau aprašytas Webster and McBratney (1989). Jis yra apskai iuojamas pagal sekan i formul :

$$AIC = -2 \ln(L) + 2m,$$

ia  $m$ - statistino modelio parametr kiekis,

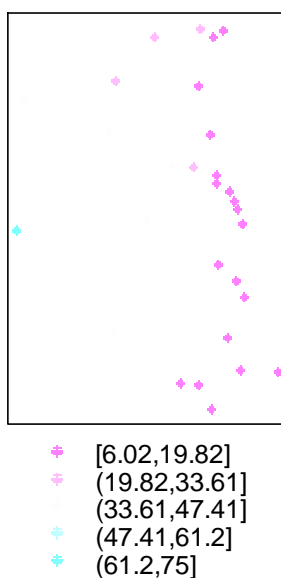
$L$  – vertinamo modelio tik tinumo funkcijos maksimali reikšm .

### III. STOCHASTINI BEI DETERMINISTINI METOD PALYGINIMAS NAUDOJANT R PROGRAMA

#### 3.1 Baltijos jūros gylio skaitmeninio reljefo sudarymas stochastiniais metodais

##### 3.1.1 Universalaus kriginio pritaikymas

Darbe bus naudojamas duomenų masyvas sudarytas iš Baltijos jūros rytinės dalies ir Kuršių mario vakarinės dalies gylio duomenų, kurie buvo išmatuoti 30 - tyje matavimo stotyse. Informacija apie 2011 metais stotyse išmatuotus gylius yra gauta ir Klaipėdoje esančio „Jūros tyrimo centro“. Atstumas tarp matavimo stoties nuo 20 - ties iki 70 - ties kilometrų. Šie duomenys bus naudojami siekiant sudaryti kuo tikslesnį skaitmeninį dugno reljefą. Iš pradžių bus atliktas skaitmeninio dugno reljefo sudarymas kriginio metodais.



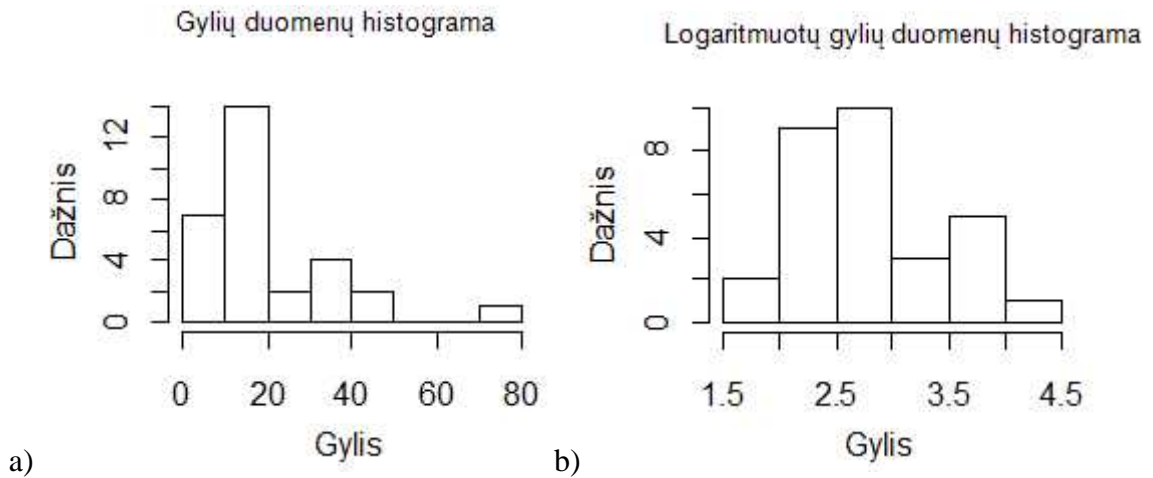
3.1 pav. Baltijos jūroje išmatuoti gyliai bei jų koordinatės ekstremumai

3.1 lentelė. Analizuojamos Baltijos jūros teritorijos koordinatės

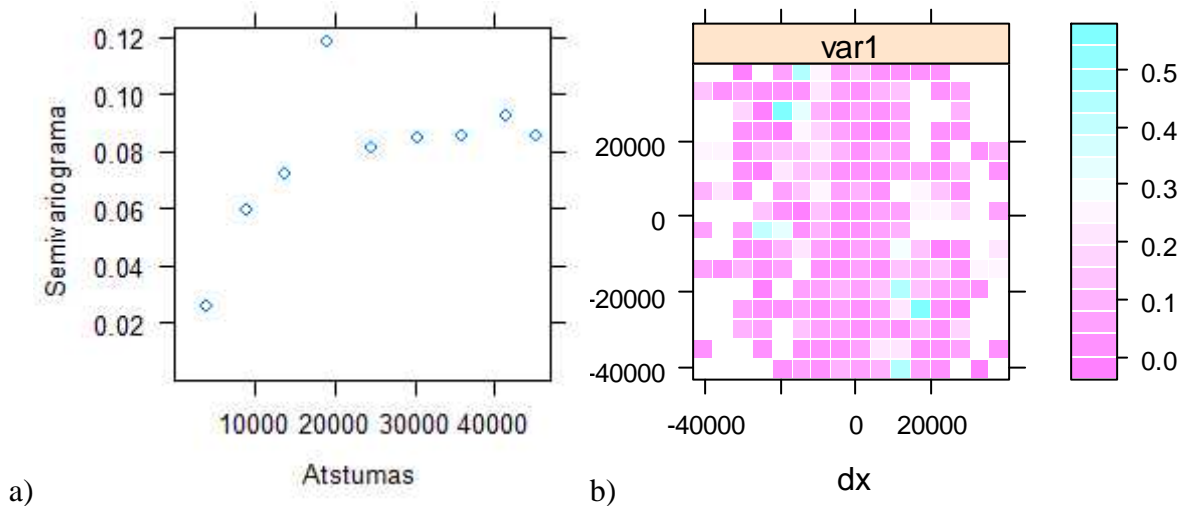
Koordinatės	Min	max
x	268984.0	328735.0
y	6128002.4	6215398.9

Prieš pradėdant darbą su duomenimis yra būtina patikrinti ar jiems yra būdingas normalinis pasiskirstymas, jeigu taip nėra duomenis yra būtina logaritmuoti. Turint dvi histogramas matome, kuriuo atveju duomenis yra tinkamesni tolesniam tyrimui. Kaip matome iš (3.2) paveikslo logaritmuotus duomenis vaizdas iš esmės pasikeitė. Akivaizdu, jog kai duomenys yra logaritmuoti jiems yra labiau būdingas normalinis pasiskirstymas. Toliau logaritmuotiems duomenims sudaroma empirinė semivariograma bei nustatoma ar duomenims būdingas trendas ir jei būdingas tai kelintos eilės (3.3 pav.). Analizuojamu atveju daroma prielaida, kad duomenims būdingas 3 – iojo laipsnio

trendas (kadangi funkcija  $g$  yra pastovum). Padaryt prielaid b tina patikrinti. Tai galima padaryti keliais skaitiniais metodais: atlikus ANOVA analiz (2.37) bei apskai iavus Akaik s informacijos kriterij (2.39).



3.2 pav. a) Logaritmuot bei b) nelogaritmuot gyli duomen histogramos



3.3 pav. a) Semivariograma, pašalinus III – ojo laipsnio trend bei b) jos žem lapis

Atlikus ANOVA analiz (3.2 lentel), gauti rezultatai patvirtino anks iau padaryt prielaid, jog modeliui analizuoti tinkamiausias yra tre iojo laipsnio trendas. Gauta F kriterijaus  $p$  reikšm yra mažiausia analizuojant kubin trend, be to ne mažiau svarbu, kad ji mažesn už reikšmingumo lygmen  $\alpha = 0,05$ . Tai reiškia, kad atmetama nulini hipotez ir teigiama, kad bent vienas regresijos koeficientas n ra nulinis (t.y gylis priklauso nuo koordina  $i$ ). Analizuojant tre iojo laipsnio trendo tinkamum gautas regresijos koeficientas 0,97 taip pat patvirtina modelio tinkamum (modelis beveik idealiai tinka duomenims prognozuoti). Remiantis Akaik s informacijos kriterijumi modeliui tinkamesnis yra taip pat tre iojo laipsnio trendas, kadangi AIC reikšm yra mažiausia (3.2 lentel).

Taigi atlikus ANOVA analiz bei skaitiniais metodais apskai iavus Akaik s informacijos kriterij nustatyta, kad modeliui analizuoti yra tinkamiausias tre iojo laipsnio trendas. Žinome, jog tiriant Baltijos j ros dugno tam tikr teritorij visai nat ralu, kad gyliai n ra išsid st kreiv s,

panašios ties pavidalu, d l vairi lank , s siauri ir t.t Taigi, nusta ius trendo laipsn tolimesniems skai iavimams b tina pateikti polinomin s lygties koeficientus  $S_i, i=1,2,\dots,10$ . (3.3 lentel ).

Toliau tiriama ar duomenims b dinga anizotropija ir jei b dinga tai kokio laipsnio.

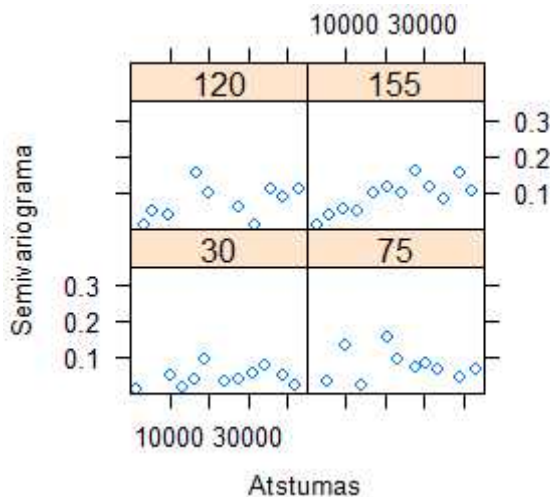
**3.2 lentel . ANOVA analiz s rezultatai, Akaik s informacijos kriterijaus (AIC) reikšm**

Trendo laipsnis	F kriterijaus reikšm	F kriterijaus $p$ reikšm	AIC	Determinacijos koeficientas $R^2$
n ra trendo	NaN	Na	165.9326	0
1	66.51861	3.6849e-11	116.5456	0.8313
2	31.21058	9.3695e-10	115.4767	0.8667
3	74.13086	2.1727e-13	77.82682	0.9709

**3.3 lentel . Tre iojo laipsnio trendo polinomin s lygties koeficientai**

$S_i$	1	2	3	4	5	6	7	8	9	10
	39.14	-42.95	-3.38	3.33	5.69	2.77	-16.99	-23.15	35.37	1.06

Remiantis (3.3) paveikslu galima daryti išvad , jog duomenims b dinga anizotropija, kadangi požymio kaita visomis kryptimis n ra vienoda. Remiantis žem lapiau sp jama, kad duomenims b dinga 30 - 35 laipsni anizotropija. Žinant anizotropijos laipsn bei nusta ius, kad duomenyse yra 3 – iojo laipsnio trendas yra braižomos kryptingos semivariogramos, nustatomas anizotropijos tipas, pagrindiniai parametrai bei empirin s semivariogramos taškams glodinamas geriausias semivariogramos modelis (Gauso, eksponentinis arba sferinis) .



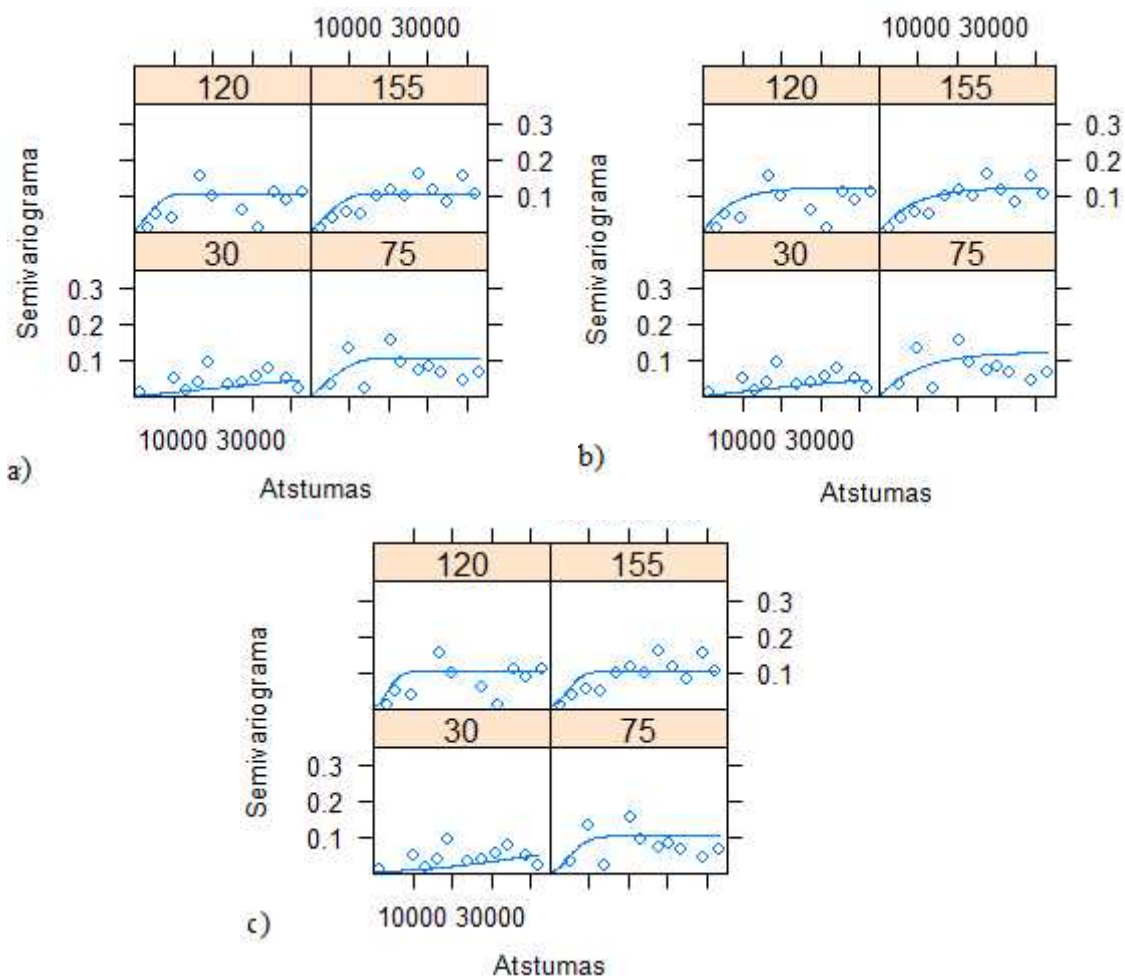
**3.4 pav. Kryptingos semivariogramos**

Iš krypting semivariogram grafiko matome, jog semivariogram su maksimaliu ir minimaliu plo iu slenks iai yra skirtingi tod l galime daryti išvad , jog turime zonin s anizotropijos atvej . Nusta ius trend bei anizotropij (jos laipsn bei tip ), pagal krypting semivariogram grafikus (3.4) nustatomi parametrai (grynuolis, slenkstis, plotis) bei empirin s semivariogramos taškams priglodinamas vienas geriausi semivariogramos modeli naudojantis

formul mis (1.1) – (1.3), o modelio tinkamumas gali būti nustatomas skaitini metod pagalba (nagrin jamu atveju kvadratinė nuokrypi suma (SS) apskaičiuojama naudojantis 1.4 formule).

**3.4 lentelė . Parametriniai semivariogramo modelių formulės**

	30°	120°
Sferinis	$\chi(d; H) = \begin{cases} 0,05 \left[ \frac{3\ d\ }{2 \cdot 30000} - \frac{1}{2} \left( \frac{\ d\ }{30000} \right)^3 \right], & 0 \leq \ d\  \leq 30000 \\ 0,05 & , 30000 < \ d\ . \end{cases}$	$\chi(d; H) = \begin{cases} 0,11 \left[ \frac{2\ d\ }{3 \cdot 18000} - \frac{1}{2} \left( \frac{\ d\ }{18000} \right)^3 \right], & 0 \leq \ d\  \leq 18000 \\ 0,11 & , 18000 < \ d\ . \end{cases}$
Eksp. (Expon.)	$\chi(d; H) = 0,06 \left[ 1 - \exp\left(-\frac{3\ d\ }{20000}\right) \right]$	$\chi(d; H) = 0,11 \left[ 1 - \exp\left(-\frac{3\ d\ }{25000}\right) \right]$
Gauso	$\chi(d; H) = 0,05 \left[ 1 - \exp\left(-3 \left( \frac{\ d\ }{18000} \right)^2 \right) \right]$	$\chi(d; H) = 0,11 \left[ 1 - \exp\left(-3 \left( \frac{\ d\ }{20000} \right)^2 \right) \right]$



**3.5 pav. Semivariogramos taškams priglodintos a) sferinis, b) eksponentinis, c) Gauso modelis**

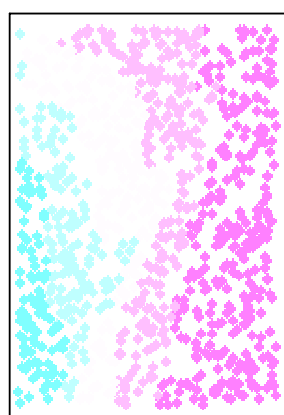
### 3.5 lentel . Priglodint semivariogram kvadratinu nuokrypi sumos

Semivariogramos modelis	SS reikšm
Gauso	2.876859e-09
Sferinis	2.875913e-09
Eksponentinis	3.077087e-09

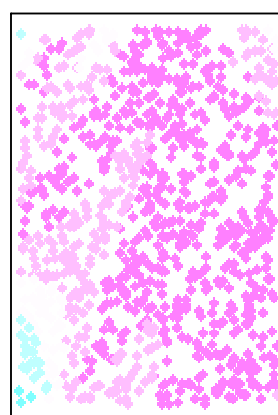
Rezultatai rodo, kad sferinis semivariogramos SS reikšm yra mažiausia (3.5 lentel ), tai reiškia geriausiai tinka jiu bus laikomas sferinis semivariogramos modelis.

Paskutinis žingsnis prieš atliekant prognoz yra tinklo (turimose teritorijos ribose) sudarymas. Ant šio tinklo ir bus prognozuojamos reikšm s naudojantis universalioju krigingu.

Taigi, išanalizav duomenis (trends tip ), nustat 50 - ties laipsni anizotropij , priglodin taškams sferinis semivariogramos model , bei sudar tinkl iš 1000 - io atsitiktinai sugeneruot koordina i teritorijos ribose, galime pasinaudojus universalioju krigingu sudaryti skaitmenin reljefo pavirši (prognoz atliekama naudojantis 2.30 formule).



+ [5.389,19.38]  
 + (19.38,33.37]  
 (33.37,47.36]  
 + (47.36,61.35]  
 + (61.35,75.34]



+ [0.002405,0.06807]  
 + (0.06807,0.1337]  
 (0.1337,0.1994]  
 + (0.1994,0.2651]  
 + (0.2651,0.3307]

### 3.6 pav. Universaliojo krigingo prognoz s rezultatai ir paklaidos

Ar duomen prognoz s yra realios sužinome palygin prognoz su turimais duomenimis bei paanalizavus variacij . Duomen variacija pagal turimas reikšmes yra pakankamai maža, ji svyruoja atsižvelgiant pradini duomen išsid stym . Didžiausia variacija yra toje teritorijoje apie kuri yra mažiau informacijos (mažiau išmatuot reikšmi ). Duomen prognoz s reikšm s iš dalies yra tinkamos, kadangi minimali ir maksimali reikšm s skiriasi visai nedaug. Tai iau toks vertinimo b das n ra patikimas ir tikslus, tod l modeliui vertinti yra naudojama kryžmin patikra. Kryžmin s patikros metodu gautos reikšm s padeda nustatyti modelio patikimum ir vertinti jo tikslum (3.11 lentel ). ia vidutin prognoz s paklaida (VPK), kuri idealiu atveju turi b ti lygi vienetui (2.34); vidutinis kvadratinis nuokrypio santykis (VKNS), kurio reikšm turi b ti artima nuliui (2.36); bei

šaknis iš vidutin s kvadratin s prognoz s paklaidos (ŠVKPS), kuri turi tur ti kuo mažesn reikšm (2.35), parodo, jog modelis n ra idealus, ta iau gana patikimas, o prognoz s rezultatai pakankamai tiksl s. Nor dami gauti patikimesnius rezultatus sekan ioje dalyje tiems patiems Baltijos j ros gyli duomenims prognozuoti bus naudojamas medianinis išlyginimo metodas

### 3.1.2 Medianinio išlyginimo krigingo pritaikymas

Norint taikyti medianinio išlyginimo kriging iš pradži reikia teritorij padalinti taisyklingas l steles (kvadratus). Perži r jus ir išanalizavus duomenis pasteb ta, kad atstumai tarp matuojam tašk yra labai skirtingi. Atstumai tarp tašk turi takos l stel s dydžio nustatyme. Taigi atsižvelgiant tašk pad t, atstumus tarp j bei j kiek (30) yra sudaromas tinklas iš 96 l steli .

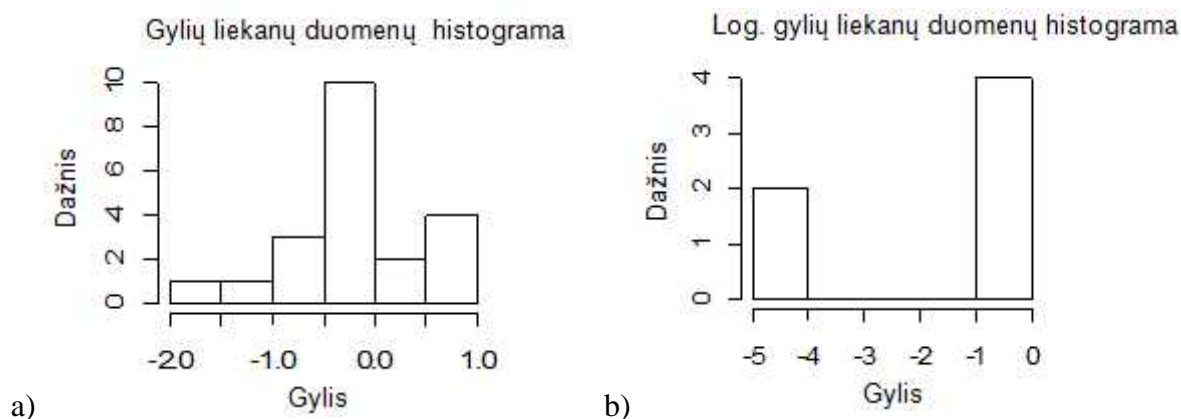
**3.6 lentel . Duomenys apie tinkl**

	X	Y
L stel s centras	272721.5	6131740.5
L stel s dydis	7477	7477
L steli skai ius	8	12

Dabar iš gardeli sudaryt tinkl apjungiamo su turimomis reikšm mis. Kiekviename gardel s narvelyje gali b ti tik po vien reikšm , tod l turime nurodyti metod , kuriuo bus apskai iuojama reikšm gardel je, jei j patenka daugiau nei vienas išmatuotas gylis (daugiau nei viena reikšm ). Galima rinktis iš trij variant , skai iuoti reikšmi mod , median arba vidurk . Kadangi analizuojamos reikšm s visos yra labai skirtingos modos nenaudosime. Mediana yra netinkama, kadangi gardel se, kuriuose yra daugiau nei viena reikšm j yra dvi ir tai tik keliais atvejais. Taigi išanalizavus duomenis matome, kad analizuojam duomen atveju geriausia yra naudoti vidurkio metod . Taip gauname tokias gardeli l steles kurias patenka po vien reikšm . Turime tinkl iš taisykling gardeli , kuriuose yra po vien reikšm , ta iau yra gardeli kurios yra tuš ios, taigi jas reikia užpildyti. Iš pradži mums reikia sudaryti matric sudaryta iš gyli , kurioje NA (yra pažymima kaip nežinoma reikšm ) žymi tuš ios gardel s reikšm . Tam sudaromas vektorius kuris turi tiek pat reikšmi kiek yra gardeli (8x12). Kiekvienai reikšmei nurodomas stulpeli kiekis 3 (toks kaip pradin se duomenyse „gyliai“). Kiekvienam stulpeliui priskiriamos koordinat s ir gyliai. Jeigu gardel tuš ia (gylis jos teritorijoje n ra išmatuotas), reikšm laikoma nežinoma ir pažymima NA. Taip gauname gyli tinkl sudaryt iš 96 gardeli , kuri kiekviena yra užpildyta. Turimas tinklas turi b ti išsaugomas kaip matrica, kadangi toliau skai iuosime eilu i bei stulpeli efektus.

Kitas žingsnis yra bendro vidurkio, eilu i bei stulpeli median efekt radimas (2.33 formul ). Ši reikšmi radimo algoritmas veikia pakaitomis šalindamas eilu i ir stulpeli medianas ir sustoja tuomet, kai absoliu i liekan sum proporcija yra mažesn už nurodyt reikšm . Iš gaut rezultat išskiriame liekanas ir jas išsaugome kaip vektori . Toliau liekanoms

priskiriamos turimos koordinat s ir taikomas ordinaraus krigin go prognozavimo metodas teigiant, kad liekanoms n ra b dingas trendas, vidurkis yra pastovus bet nežinomas.



**3.7 pav. Nelogritmuot ir logoritmuot gyli duomen histogramos atitinkamai**

Kaip matome iš liekan gyli histogram (3.7) duomenims b dingas normalinis pasiskirstymas tuo atveju, kai duomenys n ra logoritmuojami. Taigi tolimesn je analiz je bus naudojami nelogoritmuoti duomenys.

**3.7 lentel . Pagrindines liekan statistin s charakteristikos**

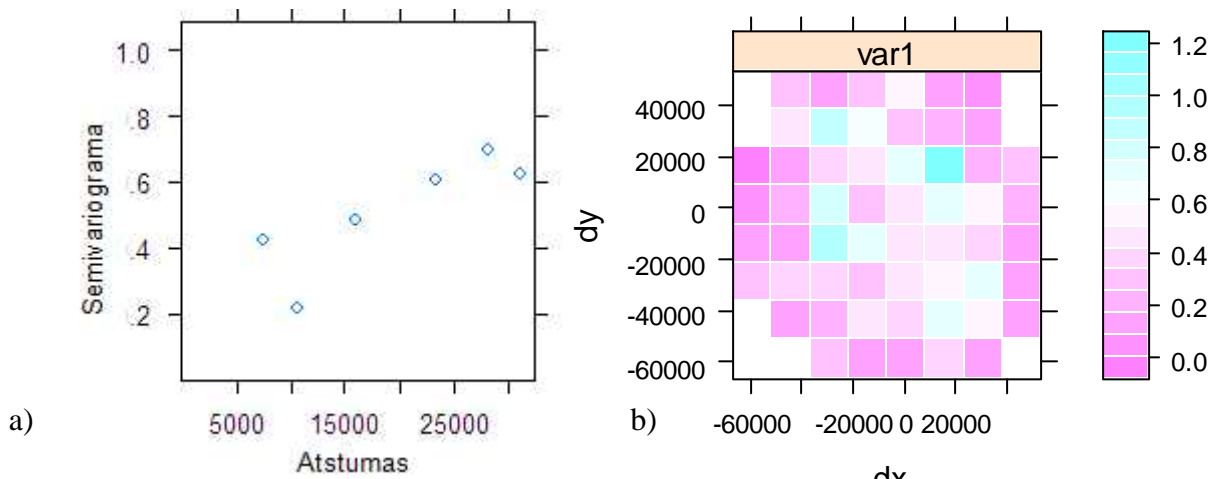
Min	1 kvartilis	Mediana	Vidurkis	3 kvartilis	max
-1.97000	-0.02832	0.00000	-0.10090	0.01479	0.96710

Sekan iame etape remiantis formule (2.21) br žiama semivariograma ir teigiama, jog liekanoms n ra b dingas trendas (ši prielaida daroma naudojantis medianinio išlyginimo krigin go teorija). Ši prielaid taip pat galima patikrinti skaitiniais metodais apskai iavus AIC kriterij (3.7 lentel ). Atliekant tolimesn analiz tirama anizotropija (tipas bei laipsnis).

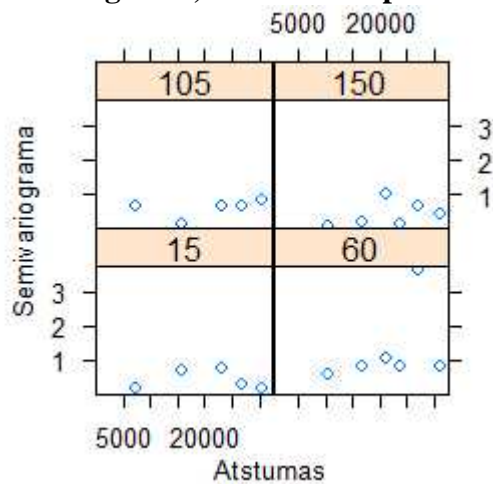
**3.8 lentel . Akaik s informacijos kriterijaus reikšm**

	N ra trendo	1 – ojo l. trendas	2 – ojo l. Trendas	3 – iojo l. Trendas
AIC	-16.98247	-13.73229	-11.07415	-5.055499
S	-0.1008784	–	–	–

Remiantis (3.8) paveikslu anizotropija yra, bet n ra akivaizdu kokio ji laipsnio. Galima daryti prielaid , jog nagrin jamu atveju egzistuoja apie 145 - 150 laipsni anizotropija. Nustaus anizotropijos laipsn br žiamos kryptingos variogramos. Nustaus variogramos pagrindinius parametrus (3.9 lentel ) priglodinamas semivariogramos modelis taikant formul mis (1.1) – (1.3) (3.10 pav.). Iš keli modeli (Gauso, eksponentinio, sferinio) išrenkamas geriausias pagal kvadratin i nuokrypi sumos formul (1.4).



3.8 pav. a) Liekan semivariograma, kai vidurkis pastovus bet nežinomas, b) jos žem lapis



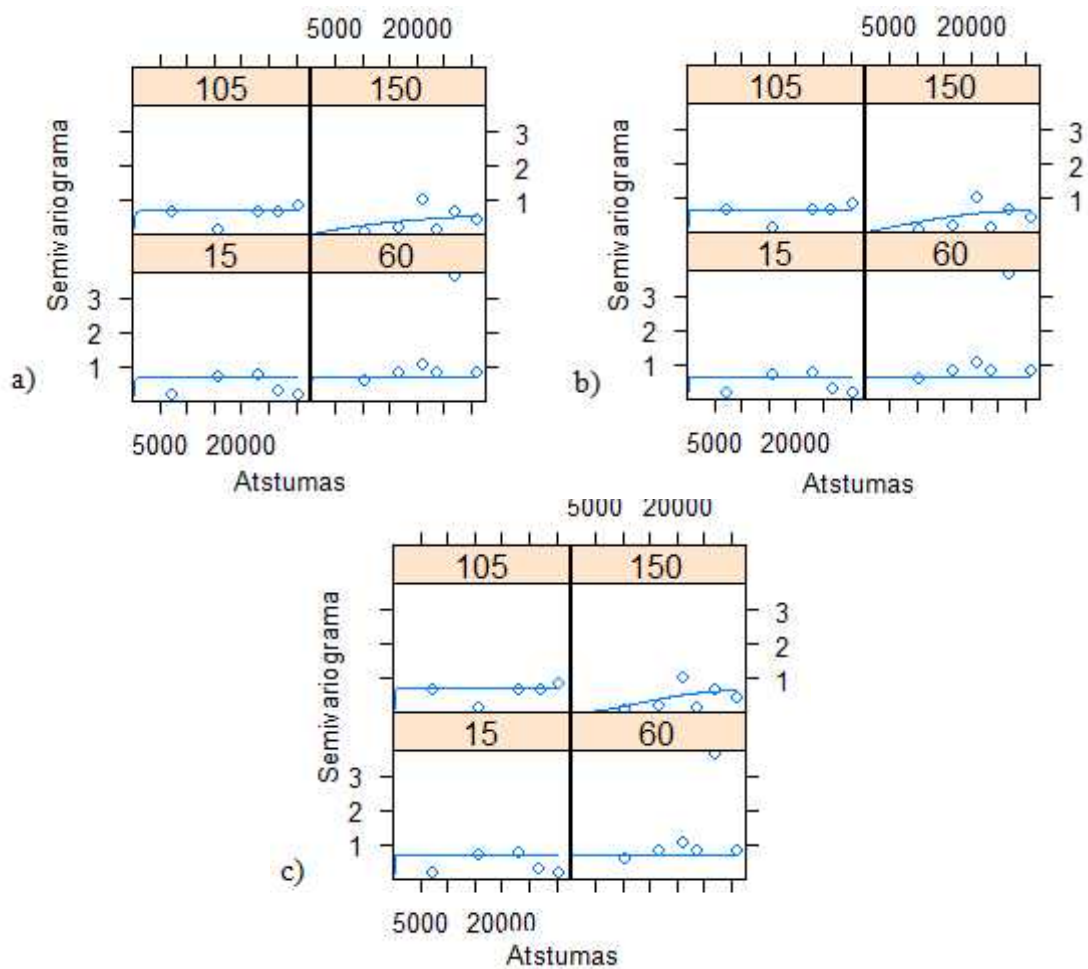
3.9 pav. Kryptingos semivariogramos liekan duomenims

Pagal (3.10) lentel galima daryti išvad, kad duomenims labiausiai tinka sferinis semivariogramos modelis, kadangi yra mažiausia SS reikšmė.

Kaip ir universaliojo krigingo panaudojimo prognozavimui paskutinis etapas prieš atliekant prognoz yra koordinuoti tinklo ant kurio atliekama prognoz sudarymas.

3.9 lentel. Parametriniai semivariogram modeli formulės

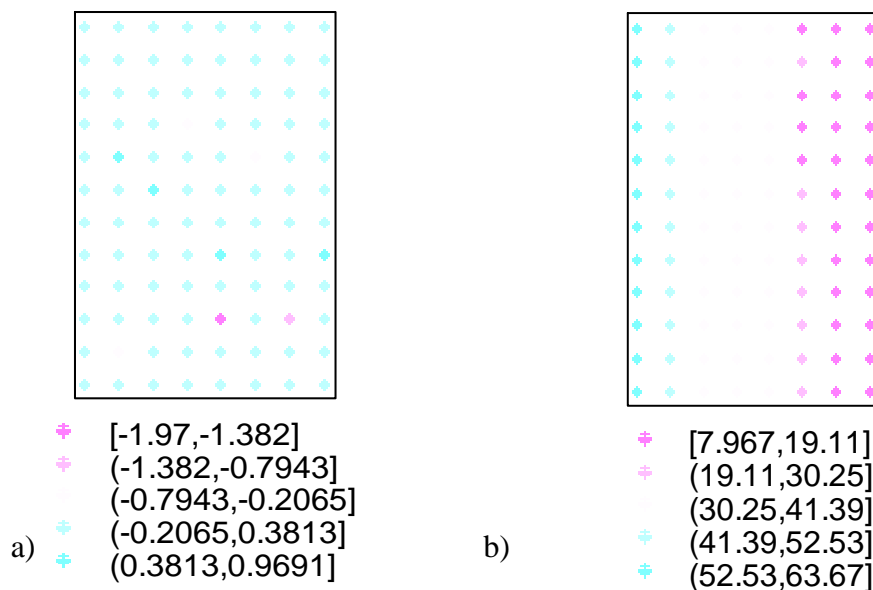
	150°	60°
Sferinis	$x(d; \Theta) = \begin{cases} 0,65 \left[ \frac{2\ d\ }{3 \cdot 30000} - \frac{1}{2} \left( \frac{\ d\ }{30000} \right)^3 \right], & 0 \leq \ d\  \leq 30000 \\ 0,65, & 30000 < \ d\ . \end{cases}$	$x(d; \Theta) = \begin{cases} 0,5 \left[ \frac{2\ d\ }{3 \cdot 25000} - \frac{1}{2} \left( \frac{\ d\ }{25000} \right)^3 \right], & 0 \leq \ d\  \leq 25000 \\ 0,5, & 25000 < \ d\ . \end{cases}$
Ekspon.	$x(d; \Theta) = 0,8 \left[ 1 - \exp\left(-\frac{3\ d\ }{25000}\right) \right]$	$x(d; \Theta) = 0,5 \left[ 1 - \exp\left(-\frac{3\ d\ }{10000}\right) \right]$
Gauso	$x(d; \Theta) = 0,7 \left[ 1 - \exp\left(-3 \left( \frac{\ d\ }{18000} \right)^2 \right) \right]$	$x(d; \Theta) = 0,5 \left[ 1 - \exp\left(-3 \left( \frac{\ d\ }{12000} \right)^2 \right) \right]$



3.10 pav. Priglodintas a) eksponentinis, b) sferinis, c) Gauso modelis

3.10 lentel . Priglodint semivariogram kvadratinu nuokrypi sumos

Semivariogramos modelis	SS reikšm
Gauso	7.269098e-08
Sferinis	6.391528e-08
Eksponentinis	9.649561e-08



3.11 pav. a) Liekan prognoz bei b) galutin gyli duomen prognoz

Atlikus liekan prognoz (remiantis (2.23) formule), būtina atkreipti dėmesį į variacijų (žr. Priede 82 psl.). Liekanų variacija atsižvelgiant reikšmes yra pakankamai didelė (variacijos reikšmės liekanoms yra beveik tokios pat kaip variacijos reikšmės patiems gyliams interpoliuojant universaliojo kriginio metodu), todėl gauti rezultatai tikslumas tiriant Baltijos jūros duomenų reljefą medianiniu išlyginimo metodu yra abejotinas, tačiau tai būtina patikrinti. Taigi, tam kad gauti skaitmeninio reljefo paviršių nagrinėjamuose reikšmių ribose prie liekan pridėdami eilutės, stulpelių medianų efektai bei bendras vidurkis (3.11 pav. b)).

Taip gavome Baltijos jūros gylių duomenų prognozę 96 taškuose naudojantis medianinio kriginio metodu. Pasinaudojus kryžminis patikros metodu galime palyginti prognozę su rezultatais tikslum ir metodų panaudojimo tinkamumą esamiems duomenims. Tam yra skaičiuojama vidutinė prognozes paklaida (ME) ir vidutinis kvadratinis nuokrypio santykis (MSRD) (2.34 – 2.36).

**3.11 lentelė . Kryžminis patikros rezultatai**

	Medianinis kriginas	Universalusis kriginas
ME	-0.83651289	0.597514
MSRD	0.6730376	0.5446689
RMSPE	5.45232	1.585214

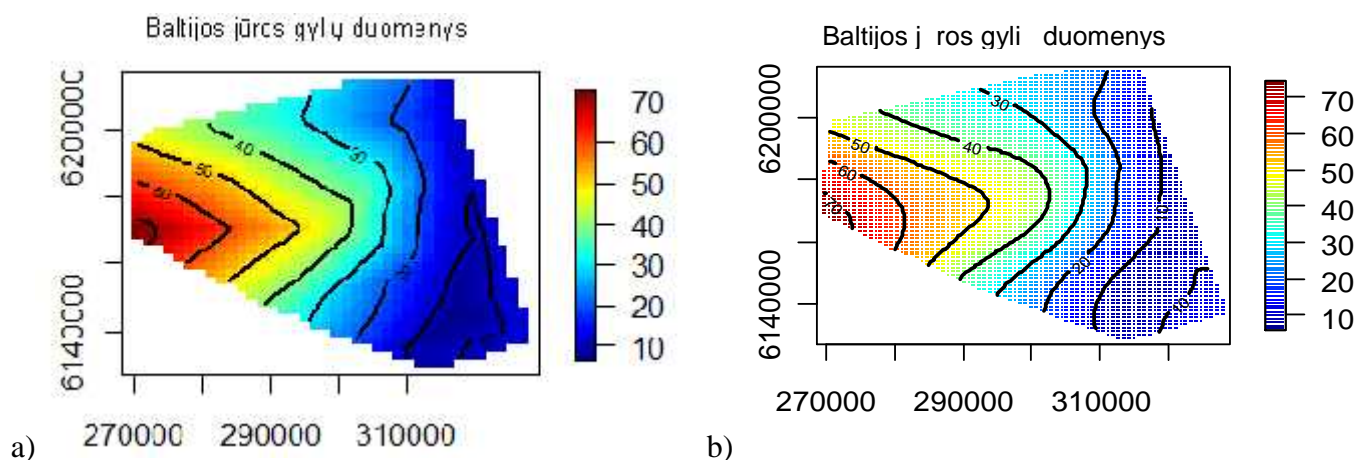
Atlikus kryžminį patikrą abiem atvejais iš (3.11) lentelės matome, jog universaliojo išlyginimo kriginio metodas yra žymiai tinkamesnis ir gauta prognozė yra tikslesnė nei medianinio kriginio metodu. Viena iš tokių kryžminis patikros rezultatų priežasčių gali būti nepakankamas duomenų kiekis atliekant medianinį kriginį. Todėl sudarintą tinklą iš 30 duomenų nagrinėjami tik 21, o vidurkis duomenų patekusių vienu laistu yra netikslus, kadangi skirtumai tarp gylių tam tikrose matavimo vietose nurodytu atstumu yra labai didelis.

Taigi gavome skaitmeninį reljefo paviršių dvejais metodais – universaliojo ir medianinio išlyginimo kriginio pagalba ir nustatėme, kuris iš jų tinkamesnis. Tačiau remiantis teorija yra teigiama, kad deterministiniais metodais gauta prognozė yra tikslesnė, todėl palyginimui sekantįjame poskyryje skaitmeninis reljefas bus sudarytas spline metodais.

### **3.2 Baltijos jūros gylių skaitmeninio reljefo sudarymas deterministiniais metodais**

Kaip jau ir buvo minėta pats paprasčiausias ir lengviausias interpoliavimo būdas yra tiesinis trikių - laiptuota interpoliacija. Jos pagrindinė savybė, kad kiekviename žingsnyje tarp eksperimentinių taškų funkcija yra konstanta. Kadangi turime realius Baltijos jūros gylių duomenis ir žinome iš bendro išsilavinimo, jog Baltijos jūros dugnas nėra lygus, šis metodas taikyti nėra naudinga bei produktyvu. Jis būtų tinkamas jei tirtume valiutos kurso kitimą. Tačiau nagrinėjami atveju yra tinkamesnis kitas tiesinis interpoliacijos tipas trikių – tiesinis interpoliacija,

kuri ieškoma priklausomybė vaizduoja laukštę. Interpoliuojama funkcija yra sudaryta iš tiesių, kurios jungiasi eksperimentiniuose taškuose – mazguose (3.13 pav.). Interpoliuojant gali būti naudojamos kelios iš pateiktų anksčiau formulė (2.9) arba (2.10).

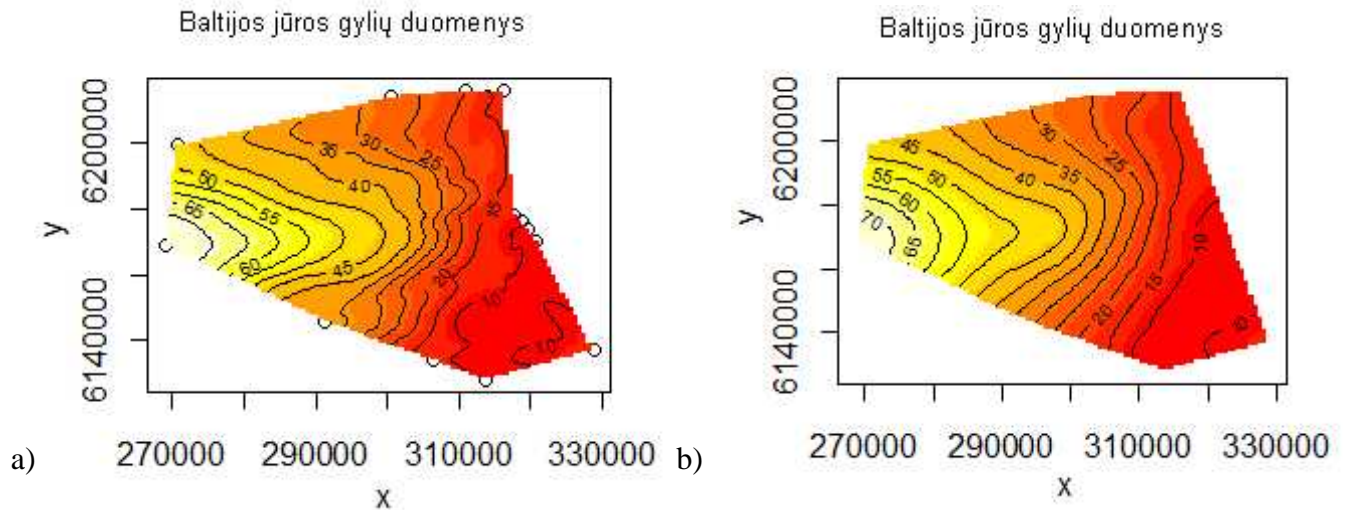


**3.12 pav. a) Tiesinė, b) kubinė interpoliacija splinais pritaikyta Baltijos jūros gyliams**

Kad rezultatas būtų patikimesnis ir geresnis yra siūloma šiuos mazgus sujungti ne laukštę, o todydžią kreivę. Geriausia tokiu atveju naudoti interpoliaciją splinais sudarytą iš polinominių funkcijų. Kaip jau yra žinoma interpoliacijos splinais pagrindinė savybė yra, kad kiekviename žingsnyje tarp eksperimentinių taškų yra vykdoma aproksimacija su tam tikra pasirinkta polinominiu priklausomybe. Tokiu būdu kiekviename žingsnyje naudojama skirtinga polinominė funkcija, o koeficientai parenkami tokie, kad žingsnio ribose (mazguose) kreivės susijungtų. Tokiu atveju dažniausiai naudojami yra trečiojo laipsnio polinomialai – kubiniai splinai, interpoliacija atliekama remiantis formulėmis (2.14) (3.12 pav.). Kaip ir yra žinoma Baltijos jūros gylis yra ne tik netolygus paviršius, bet pasitaiko ir duobės. Interpoliuojant duobes galima vadinti išskirtimais ir kad šios išskirtys nepadarytų didelės tokos galutiniam rezultatui galima naudoti kubinį Akima spliną (jis yra atsparus išskirtimams), (3.13 pav.). Tam, kad pritaikyti šiuos metodus turi būti tenkinama sąlyga, kad nepriklausomi vektoriai, nagrinėjamu atveju koordinatės, nėra kolinearūs (vektoriai yra kolinearūs, kai yra lygiagretūs). Šiai sąlygai patikrinti atlikta kolinearumo diagnostika, kuri patvirtino, kad koordinatės nėra kolinearios (3.11). Jeigu B senos indeksas yra didesnis už 30 ir abiejų vektorių variacijos proporcija didesnė nei 50 procentų vektoriai būtų kolinearūs (Belsley, D.A., 2001).

**3.12 lentelė. Koordinatės kolinearumo diagnostika**

	B senos indeksas	Variacijos proporcija	
		X	Y
1	1.000	0.371	0.371
2	1.301	0.629	0.629



3.13 pav. a) Kubinis Akima, b) kubinis B - splainas pritaikytas Baltijos jūros gyliams prognozuoti.

Taip pat paviršiaus interpoliavimui yra naudojami ir B – splainai. Jų laipsnis yra parenkamas priklausomai nuo tiriamos situacijos. Jie naudojami kreivi braižymui pagal atraminius taškus. Dažniausiai naudojamas yra kubinis B – splainas, todėl jis ir bus pritaikytas Baltijos jūros duomenų analizei (3.15 pav.). Šio splaino vienas iš privalumų yra tas, kad dirbant su labai dideliu kiekiu duomenų jis reikalauja mažiausiai kompiuterio atminties ir laiko lyginant su kitomis interpoliavimo funkcijomis.

Atlikus Baltijos jūros dugno reljefo prognozę keliais deterministiniais metodais: tiesiniu, kubiniu, kubiniu Akima bei kubiniu B – splainu tam, kad pasirinkti tinkamiausi bei tiksliausi metodai, būtina juos palyginti (nustatyti kiekvieno metodo tikslumą). Nagrinėjant atvejų modelio tinkamumą padės nustatyti anova analizės metu paskaičiuota  $R^2$  reikšmė.

3.13 lentelė. Modelio tinkamumo reikšmės

Splaino tipas	$R^2$ reikšmė
Tiesinis	0,80
Kubinis	0,97
Kubinis Akima	0,79
Kubinis B - splainas	0,94

Anova analizės metu apskaičiuota modelio tinkamumo duomenims reikšmė parodo, kad mažiausiai tinkantys modeliai yra tiesinio ir kubinio Akima splaino. Tai reiškia, kad interpoliuojant Akima splainu gautume labiausiai nuo realios situacijos nutolusias reikšmes. To priežastis yra greičiausiai ta, kad Baltijos jūra turi daug duobių ir anomalijų, kurias ignoruoti ir laikyti klaidomis

yra netikslinga. Tiesinio splaino modelis laikomas taip pat ne pačiu tiksliausiu. Tai galima paaiškinti tuo, kad naudojantis tiesiniu splainu gaunamas staigus reikšmių pokytis visuose tiriamuose mazguose (aštrūs kampai interpoliacijos mazguose). Du metodai, kurie anota ANOVA analizės (3.10 lentelė) yra tinkamiausi atlikti analizuojamų reikšmių interpoliacijai yra kubinis splainas ir kubinis B – splainas. Šių modelių rezultatų panašum galima taip pat pamatyti paanalizavus (3.13 – 3.14) paveikslus. Šių reikšmių pasiskirstymas ir interpoliuota teritorija taip pat pabrėžia modelių prognozių panašumą.

Taigi atlikus Baltijos jūros dugno reljefo analizę skaitiniais metodais, atlikus ANOVA analizę (F kriterijumi), bei apskaičiavus AIC reikšmę yra pasirinktas vienas stochastinis (universalus krigingas) vienas deterministinis metodas (kubinis splainas), kurie pateiks tiksliausius rezultatus interpoliuojant turimą teritoriją. Naudojantis darbe atliktais tyrimais (3.2, 3.11 lentelėmis) iš šių dviejų metodų tikslesnis bei patikimesnis yra antrasis.

## IŠVADOS

Taigi kaip yra žinoma pagrindin žemės reljefo duomen atvaizdavimo priemon yra skaitmeniniai reljefo modeliai. Šiuo atveju nagrin jamas paviršius yra Baltijos jūros dugnas. Analizuojamos gylio reikšmės, kurios pasiskirs iusios teritorijoje su atitinkamomis koordinatėmis. Pagrindinis darbe naudojamas paviršiaus atk rimo metodas duomen interpoliacija. Šiais laikais yra žinoma dešimtis interpoliacijos būdų, tačiau darbe naudojami keli populiariausi bei dažniausiai naudojami - krigingas bei interpoliacija splainais. Tam, kad atlikta interpoliacija būtų naudinga, patikima bei kuo manoma tiksli darbe yra gyvendinti šie uždaviniai:

- Darbe aprašyti pagrindiniai stochastiniai bei deterministiniai metodai: krigingas, splainas. Išsiaiškinta, kad krigingas – metodas, kurio pagrind sudaro matematinės statistikos metod panaudojimas. Jo realizacijoje naudojama regionalizuoto kintamojo idėja, o paviršius analizuojamas keli nepriklausom kintamųjų pavidale (trendas, nuokrypi kintamasis). Pagrindinis splain savyb yra ta, kad jie gali aprašyti sudingus paviršius pasinaudojant neaukšto laipsnio polinomais. Taip yra todėl, kad naudojant interpoliaciją splainais visa teritorija yra padalinama nedidelius tarpusavyje nepersikertančius plotus. Aproximacija polinomais yra vykdoma atskirai kiekvienam plotui, o vėliau sudaroma bendra funkcija, kuri apjungia vis teritoriją.
- Pateiktos bei išanalizuotos ordinarus, paprastojo, universalus, medianinio krigingo ir tiesinio, kvadratinio, kubinio, kubinio Akima ir kubinio B splaino matematinės interpretacijos.
- Pasinaudojus R programos paketais gstat, fields, Akima, mba, spatial atlikta Baltijos jūros gylio prognozė. Atlikus anova analizę, bei apskaičiavus Akaikės informacijos kriterijų nustatyta, kad gylio duomenims būdingas trečiojo laipsnio trendas, todėl buvo pritaikytas universalus krigingo bei medianinio išlyginimo krigingo metodai. Taip pat atlikus kolinearumo diagnostiką ir patvirtinus, kad koordinatės nėra kolinearios skaitmeninis reljefo paviršius sudarytas tiesinio, kubinio, kubinio Akima bei kubinio B – splaino metodais.
- Palygintos krigingo bei splaino interpoliacijos metodais gautos prognozės. Metodų palyginimui, tikslumui bei tinkamumo duomenims vertinimui atlikta anova analizė bei pritaikytas kryžminis patikros metodas. Rezultatai parodė, kad iš vis gylio duomenų prognozei taikytas metodas geriausi ir tiksliausi yra kubinis splainas bei universalus krigingas, kadangi abejais atvejais gauta modelio tinkamumo duomenims reikšmė  $R^2 = 0,97$ .

Taigi atlikus visus darbe iškeltus uždavinius, išanalizavus metodus bei tinkamai juos pritaikius duomenims buvo gauta prognozė šešiais aukščiau išvardintais metodais. Anova analizės pagalba (remiantis  $R^2$  reikšme) išrinkti keli tinkamiausi Baltijos jūros dugniui interpoliuoti metodai: universalus krigingas (teigiant, kad duomenims būdingas trečiojo laipsnio trendas) bei kubinis splainas. Šių metodų tinkamumo reikšmės yra praktiškai vienodos, todėl galime patvirtinti darbo pradžioje iškeltą hipotezę, kad sudarintą skaitmeninį Baltijos jūros dugno paviršių, turint taisyklingai išsidėstiusius duomenis, rezultatų tikslumui metodo, stochastinio ar deterministinio (kubinio splaino ar universalaus kringingo), pasirinkimas taktinis.

## LITERATŪROS RAŠAS

1. Kumetainien A. 2006. Skaitmeninio reljefo modelio sudarymas skirtingais geostatistiniais reljefo modeliavimo metodais. VGTU
2. . . . . 1990.
3. . . . . 1980.
- 4 . J. Sinclair, Garston H. Blackwell . 2007: Applied mineral inventory estimation. UK
5. Albert J., Nilson E., 1972. Spline teory.
6. Azari, A.S., Müller, H.G. 1992. Preaveraged localized orthogonal polynomial estimators for surface smoothing and partial differentiation. J. American Statistical Association 87, 1005-1017.
7. Alastair J., Garston H., 2002 . Applied mineral inventory estimation. Cambridge
8. Bernard A., Megrey, M. 2009. Computers in Fisheries Research.
- 9 . Blyth A., Cake D., Andresen M.,Gienko G., Govorov M. 2008. Geografinės informacijos infrastruktūros nuotolinis mokymas: erdvinis analizė ir modeliavimas. Mokomoji knyga
10. Chilès J.P., Delfiner P. 1999. Geostatistics modeling spatial uncertainty.
11. Geoffrey M. 1994. American Statistical Association: Kriging and Splines: An Empirical Comparison of Their Predictive Performance in Some Applications.
12. Hutchinson, M.F., and P.E. Gessler. 1994. Splines - more than just a smooth interpolator. Wageningen, Netherlands. Geoderma .45-67 p.
13. Wiley M., Sons. 2007. Geostatistics for environmental scientist.
14. Johnston K., Ver Hoef J. M., Krivoruchko K., Lucas N. 2001. Using ArcGIS Geostatistical Analyst. Redlands CA, ESRI Press. 300 p.
15. Journel, A. G., Huijbregts, Ch. J. 1978. Mining Geostatistics . Academic
16. Krige D., 1951. A statistical approach to some mine valuation and allied problems on the Witwatersrand. Unpublished Master's Thesis, University ofthe Witwatersrand.
17. O'Neill C.,2002. Cubic spline.
18. Sankar V.P., Ferrary L.A., 1998. B – spline simple algoritms and architectures for interpolation.
19. Stanford Exploration Project. 2000. B – spline interpolation.
20. Voltz M .,Webster R. 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Science 31: 505-524.
21. Wahba, G. 1990. *Spline Models for Observation Data*. Society for Industrial and Applied Mathematics.
22. Webster, R. and Oliver, M.A., 1990: Statistical methods in soil and land resource survey.

23. Sky McKinley and Megan Levine. 1998. Cubic Spline Interpolation. [interaktyvus] Prieiga per internet <http://online.redwoods.cc.ca.us/instruct/darnold/laproj/Fall98/SkyMeg/proj.pdf> [ži r ta 2010 gruodžio 13d.]
24. Wikipedia. 2010. Spline. Prieiga per internet [interaktyvus] [http://en.wikipedia.org/wiki/Spline\\_\(mathematics\)](http://en.wikipedia.org/wiki/Spline_(mathematics)) [ži r ta gruodžio 10d.]
25. ALGLIB.2010.Spline interpoliacija. [interaktyvus] Prieiga per internet <http://www.alglib.net/interpolation/spline3.php> [ži r ta 2010 gruodžio 10d.]
- 26.VDU.2009. Grafika ir vizualizavimas [http://fcis.vdu.lt/~gintare.ceidaite/grafika/Dien/Paskaitos/paskaita\\_17\\_kreives.pdf](http://fcis.vdu.lt/~gintare.ceidaite/grafika/Dien/Paskaitos/paskaita_17_kreives.pdf)
27. Kriging. Prieiga prie interneto. <http://gbi.agrsci.dk/statistics/courses/JBS-Geostatistics-2003/Day3/Lecture1/D3L1-handsout.pdf>
28. Splainai. Prieiga per internet . [http://reslib.com/book/Methodi\\_splajn\\_funkcij\\_\\_Zavjyalov\\_Yu\\_S\\_/11](http://reslib.com/book/Methodi_splajn_funkcij__Zavjyalov_Yu_S_/11)
29. Neuman, S. P. and Jacobson, E. A. 1984. Analysis of nonintrinsic spatial variation by residual kriging with application to regional groundwater levels. *Journal of the International Association for Mathematical Geology*, 16, 499–521.
30. Cressie, N. 1993. *Statistics for Spatial Data*, Wiley & Sons, New York, p. 156.
- 31.Johnston, K., Ver Hoef, J.M., Lukas, N. 2001. *Using ArcGIS Geostatistica Analyst*. Redlands CA, ESRI Press, psl 300.
- 32.Belsley, D.A (2001). *Conditioning diagnostics, collinearity and weak data in regression*. New York: John Wiley & Sons.

