

C CONSTUD

Tarkvarasüsteemi Constud kasutamissoetus

Autoriõigus: Kalle Remm ja Tiiu Kelviste

Kujundus: Tiiu Kelviste

Retsensendid:

Tiit Nilson, Olavi Kärner (Tartu Observatoorium),

Tõnu Möls (Eesti Maaülikooli Limnoloogiakeskus, Tartu Ülikooli matemaatika-informaatikateaduskond)

Reprodutseerimisõigus: Tartu Ülikool, Kalle Remm ja Tiiu Kelviste

Publitseerimine DSpace hoidlas: Tartu Ülikooli Raamatukogu

<http://dspace.utlib.ee/dspace>

Raamatu väljaandmist on toetanud Tartu Ülikooli geograafia osakond, sealhulgas Haridus ja Teadusministeeriumi sihtfinantseeritava teema SF0180049s09 vahenditest.

Constud veebileht: <http://www.geo.ut.ee/CONSTUD>

Constud versioon mai 2011 algkoodi autor: Kalle Remm

Constud algkoodi varalised õigused: Tartu Ülikool

ISBN: 978-9985-4-0666-3

Tartu Ülikooli geograafia osakond, geoinformaatika ja kartograafia õppetool,
Vanemuise 46, 51014

Tartu 2011

TARKVARASÜSTEEMI CONSTUD KASUTAMISÕPETUS

Kalle Remm ja Tiiu Kelviste

Tartu Ülikool, geoinformaatika ja kartograafia õppetool

Tartu 2011

SISUKORD

Eessõna	6
1. Tutvustus	7
1.1. Kasutamise õpetus	9
1.2. Andmebaasi ülesehitus	11
2. Andmestruktuuri loomine	16
2.1. Andmebaas	17
2.2. Andmetabelid	18
2.2.1. Andmetabelite loomine Constud abil	18
3. Mustriindeksite arvutamine	28
3.1. Indeksid andmetabelisse	30
3.1.1. Valimi kasutamine	31
3.1.2. Indeksite arvutamine eraldiste piires	33
3.1.3. Õmblusteta andmekihid	35
3.1.4. Indeksid prognoositabelitesse	36
3.2. Indeksid rasterfaili	38
3.3. Indeksid kaheбайдise täisarvu vormingus andmekihist	41
4. Kaalude sobitamine	44
4.1. Tehisõpe ehk kaalude automatiseeritud sobitamine programmis Constud	45
4.1.1. Sarnasuse arvutamine	48
4.1.2. Hinnangu arvutamine	49
4.1.3. Tõesuskriteerium	50
4.2. Constud õppe suvandid	50
4.3. Funktsioontunnuse eelklassifikaator	52
4.4. Kaheväärtuseline muutuja	53
4.5. Multinomiaalne muutuja	57
4.6. Numbriline muutuja	59
4.7. Üks multinomiaalse muutuja üksus	61
4.8. Mitte-ruumiline muutuja	65
4.9. Mitmemõõtmeline muutuja	67
5. Teadmiste kontroll	69
5.1. Logitabelite sirvimine	70
5.2. Kontrollvastavuse arvutus	71
5.3. Näidiste uuesti valik	75
6. Hinnangute arvutus constud andmebaasi	77
6.1. Kaheväärtuseline muutuja	78
6.1.1. Ruumiline kaheväärtuseline muutuja	79
6.1.2. Mitteruumiline muutuja	80
6.2. Multinomiaalne muutuja	81
6.3. Pidev muutuja	82
6.4. Multinomiaalse muutuja eraldi õpitud üksused	83

6.5. Sarnasus etteantud üksusega	84
7. Hinnangulised kaardid	88
7.1. Interpoleerimine	89
7.2. Üksikmuutuja kaart	91
7.2.1. Kaheväärtuseline muutuja	92
7.2.2. Multinomiaalne muutuja	95
7.2.3. Numbriline muutuja	97
7.2.4. Nominaalse muutuja üksus	99
7.3. Komplekshinnang	101
7.4. Asendustunnused	105
Lõppsõna	108
Viidatud kirjandus	109
Lisad	110
Lisa 1. Mõisted	111
Lisa 2. Constud eestikeelse versiooni andmetabelid	115
Lisa 3. Mustriindeksid süsteemis Constud	121
3.1. Nominaalsest andmekihist arvutatavad indeksid	121
3.2. Numbrilisest andmekihist arvutatavad indeksid	122
Lisa 4. Küsimuste vastused ja selgitused	124

EESSÕNA

Constud lähtekoodi on kirjutanud Kalle Remm aastatel 2002–2011. Madli Linder, Eerik Absalon, Hendrik Proosa, Tanel Tamm, Tiiu Kelviste, Allan Rajavee ja Joonas Remm abistasid süsteemi katsetamisel ja tehnilise dokumentatsiooni kirjutamisel. **Constud** lähtekood on Tartu Ülikooli omand.

Constud süsteemi kirjeldus on publitseeritud rahvusvahelise levikuga väljaannetes, nagu *Ecological Modelling* (Remm, 2004), *Artificial Intelligence in Medicine* (Remm ja Remm, 2008), *International Journal of Applied Earth Observation and Geoinformation* (Tamm ja Remm, 2009) ja *Health & Place* (Remm ja Remm, 2010).

Autorid avaldavad tänu professor Ülo Manderile selle juhendi koostamisel abiks olnud finantstoe leidmise eest, Jüri Roosaarele ja Kiira Mõisjale abi eest mõistete määratlemisel, süsteemi ja juhendi eelpool mainitud katsetajatele, 2011. aasta kevadsemestril õppeaines Kaugseire 2 osalenud tudengitele parandusettepanekute eest ning neile, kes stimuleerisid selle juhendi kirjutamist väites, et süsteemi **Constud** kasutama õppimiseks kulub vähemalt viis aastat. Loodame, et see juhend aitab lühema ajaga toime tulla.

Süsteemi loojal oleks hea meel näha süsteemi kasutajate hulga kasvu, kuid on ettevaatlik süsteemi soovitamisel juhukasutajatele, sest **Constud** toimib vaid koos andmebaasiga, milles on kindlalt ettenähtud struktuuriga tabelid. Selle tarkvarasüsteemi jaoks vajaliku andmebaasistruktuuri mõistmine ja loomine on süsteemi õppimisel ja kasutamisel esmatähtis.

Kes õpib, see oskab!

Autorid

TUTVUSTUS 1

CON  STUD

Constud on MSWindows süsteemis töötav tarkvara, mille nimi tuleneb ingliskeelsetest sõnadest *Continuous Study*. Constud tarkvara on testitud op-süsteemides Windows XP, Windows Vista ja Windows 7. Süsteemi kolm põhilist kasutusvaldkonda on järgmised:

- 1) ruumimustri kirjeldavate indeksite arvutamine rasterkujul andmekihtidest andmebaasi tabelitesse või rasterfaili;
- 2) tehiseõpe ehk tunnuste ja vaatluste komplekti iteratiivne valik ning kaalude sobitamine õpetusandmetele;
- 3) nominaalsete ja arvuliste muutujate sarnasusele tuginev prognoosimine andmebaasi tabelitesse või rasterfaili.

Tavapäraselt arvutatakse prognoosid mudelist või saadakse eksperthinnanguna. Constud on tarkvarasüsteem sarnasusele tuginevate ehk näidistele tuginevate järelduste arvutamiseks. See tähendab, et funktsioontunnuse ehk prognoositava muutuja väärtus tuletatakse sama muutuja väärtustest kõige sarnasematel näidistel. Sarnasusele tugineva järeldamise puhul teoreetilist mudelit ei kasutata, mis teeb sarnasuse järgi järeldamise lihtsamaks võrreldes mudelite abil prognoosimisega. Ainsad reeglid, mida on sarnasusele tugineva järeldamise jaoks tarvis, on sarnasuse arvutamise viis ja õpetusandmete või varasemate teadmiste hulgast kõige sarnasemate näidiste otsimise meetod. Erinevalt mudelite kasutamisest, ei kaasata näidistele tuginevasse prognoosisüsteemi mudeli valija subjektiivset arusaama seostest tunnuste vahel.

Constud süsteemi potentsiaalsed kasutajad on isikud, 1) keda huvitavad andmekaevandamise ja mustrituvastamise meetodid, 2) kes vajavad tarkvara hinnanguliste kaartide arvutamiseks.

Constud on üsna universaalne ja mitmete võimalustega sarnasuse järgi hindamise süsteem, mille eelisteks on esiteks sarnasusele tuginev järeldamine, mis tagab hinnangute arvutamise igat tüüpi andmetest. Teiseks on Constud hinnangulise kaardistamise etappe integreeriv lahendus — samasse tarkvarasüsteemi on ühendatud kohta kirjeldavate tunnuste (kohatunnuste) arvutamine, prognoosiva süsteemi sobitamine õpetusandmetele ja hinnanguliste kaartide genereerimine. Constud on lähedane andmekaevandamise meetoditele, mõlema puhul on kasutamise eesmärk eelkõige praktiline — parima võimaliku hinnangu saamine suurest paljude tunnustega andmete hulgast, mitte statistiliste hüpoteeside tõestamine. Constud ei arvuta tõenäosusi ega statistilist olulisust.

Constud süsteemi on ühendatud kohatunnuste
 arvutamine, prognoosiva süsteemi sobitamine
 õpetusandmetele ja hinnanguliste kaartide
 genereerimine

Süsteem koosneb kolmest osast :

- 1) seletavate tunnuste, **eelklassifikaatorite** ja interpoleerimiseraldiste (vt peatükk 7) rasterkujul **andmekihid** (kui kasutatakse ruumiandmeid);
- 2) **andmebaas** vaatluste, tehisõppe parameetrite, tehisõppe tulemuste ja tabelisse prognoositud väärtuste hoidmiseks;
- 3) **tarkvara Constud**.

Constud süsteemi osad :

andmekihid
andmebaas
tarkvara Constud

Constud andmekihid peavad olema kokkupakkimata, päiseta binaarsete rasterfailidega (**Idrisi32** vormingus rst-failid). **Vektorkujul** andmekihid tuleb rasteriseerida, muus vormingus rasterkihid teisendada. Samal kaardilehel peavad olema samad servakoordinaadid kõigis andmekihtides. Ühesuuruste ruudukujuliste kaardilehtede korral ei ole kõigi servade ristkoordinaate tarvis, piisab x ja y miinimumist ja kaardilehe serva pikkusest. Tarkvarapaketi **Idrisi** uusimat versiooni saab osta aadressilt: <http://www.clarklabs.org/>.










Juhendis kasutatud spetsiifilisemad mõisteid on seletatud lisas 1. Juhendis olevad SQL käsud, **Constud** süsteemi tehniline kirjeldus, käivitamise ja installeerimise viidad, publikatsioonide loetelu ning põhiliste funktsioonide lähtekood on **Constud** veebilehel <http://www.geo.ut.ee/CONSTUD>. Veebilehelt saab teatada selle juhendi ja **Constud** süsteemi vigadest. Samas avaldatakse ka selles juhendis ilmsikstulnud vead ja süsteemi muudatused.

Installeeritud **Constud** kontrollib iga käivituse eel interneti kaudu, kas serveris on tarkvara uuendus. Kui on, siis küsib, kas uuendus installeerida. Programmi nimi võib sisaldada versiooni numbrit, näiteks kujul **Constud 2**.

1.1. KASUTAMISE ÕPETUS

Süsteemi **kasutamise näidisülesanded on rasvases kirjas ja nummerdatud**. Numeratsioon on peatükke läbiv, sest reeglina ei saa järgnevat sammu teha enne, kui kõik eelnevad sammud on edukalt läbitud. Tarkvarasüsteemi **Constud** kasutus peaks üldjuhul algama katsete ja kasutatavate tunnuste kavandamisest ning andmekihtide ettevalmistamisest. **Constud** kasutajalt eeldatakse **Access** või **SQL Server** andmebaasiga töötamise kogemust ning SQL keele, geoinformaatika ja andmetöötuse põhitõdede mõistmist vähemalt algtasemel.

Ülesannete juures on järgmise tähendusega ikoonid:

-  — Constud programmi või arvutuse käivitus.
-  — Constud programmi **suvandite** valik.
-  — nupule vajutus järgmise akna avamiseks.
-  — tabeli või andmebaasivaate loomine või kopeerimine.
-  — tabeli või andmebaasivaate ehituse või selles olevate andmete muutmine, välja arvatud uuritava muutuja valik.
-  — muutujate sisse- või väljalülitus väljalt *[FTUNNUSED].[arvutada]*.
-  — andmete andmetabelisse või failide kasutaja arvutisse lisamine.
-  — suvandite ja andmete kontroll.
-  — tulemustega tutvumine ja nende mõtestamine, iseseisev katsetamine.

Andmebaasi muutmise sammud on valdavalt esitatud SQL keeles päringutena. **Access** ja **SQL Server** pakuvad vahendeid samade tööoperatsioonide tegemiseks tabelite disainvaates ja tabelivaates, interaktiivselt hiire ja klaviatuuri abil. Loomulikult võib neid vahendeid kasutada. Juhendi tekstis on eelistatud tööoperatsioonide SQL keele kuju, kuna see sõltub vähem andmebaasihalduri kasutajaliidese **suvanditest** ning teksti kujul päringuid on lihtsam muu tekstiga lõimida. Suur- ja väiketähtede kasutamine **andmebaasiobjektide** (selles juhendis tabelid ja päringud) nimedes ja SQL keele sõnades on vaba — SQL ei ole tõstetundlik.

Tekstilaadiga on välja toodud lisas 1 selgitatud **mõisted** ning *andmebaaside* ja *andmebaasiobjektide* nimed. Andmebaasiobjektide nimed, mis koosnevad tabeli nimest ja välja nimest ning enamik andmebaasiobjektide nimesid SQL-keelsetes lausetes, on nurksulgudes (näiteks *[FTUNNUSED].[FID]*). Tänu sellele võib andmebaasiobjekti nimi sisaldada reserveeritud sõnu ja sümboleid. Mõistele klikates kuvatakse mõiste tähendus, tagasi mõiste juurde tekstis saab Alt+ „vasak nool“ klahvikombinatsiooniga. Samuti on taaskuvatavad tekstis viidatud joonised ja näidisülesanded.

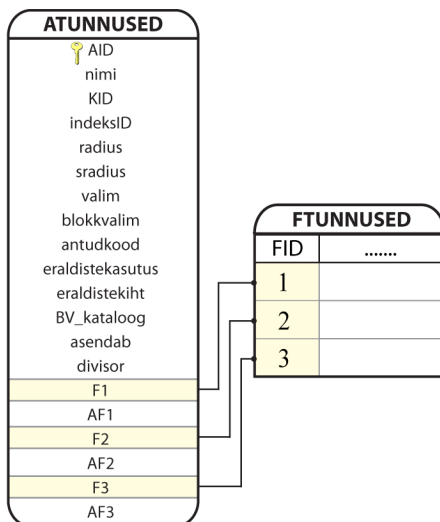
Ülesannete lahendamiseks vajalikud tabelandmed on näidisandmete andmebaasis *Constud_6ppe_andmed.mdb*. Seda faili saab alla laadida **Constud** veebilehelt.

Näidisharjutustes kasutatavad rasterfailid on Tartu Ülikooli Ökoloogia ja Maateaduste Instituudi (TÜ ÖMI) serveris, mille FTP serveri aadress ja autentimislause on **Constud** programmi sisse kirjutatud. Väljaspool TÜ arvutivõrku töötades tuleks kasutada VPN ühendust (vt <http://arvutiabi.ut.ee/pages/viewpage.action?pageId=3309798>). Teave ÖMI serveris oleva digitaalsete kaardi- ja kaugeireandmete arhiivi kohta on veebilehel <http://digiarhiiv.ut.ee>. Kasutajad väljaspool Tartu Ülikooli arvutivõrku peaksid andmekihid ise looma või hankima. Andmekihtide ettevalmistamise juhised on süsteemi tehnilises kirjelduses (*Remm ja Linder, 2008*).

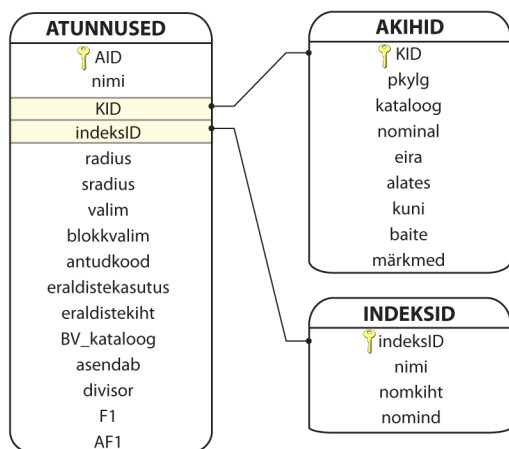
Iga peatüki lõpus olevate küsimuste vastused on lisas 4.

1.2. ANDMEBAASI ÜLESEHITUS

Constud on loodud põhiliselt prognoosiülesande täitmiseks, see tähendab **funktsioontunnuse** ehk uuritava muutuja väärtuse prognoosimiseks **seletavate tunnuste** ehk **argumenttunnuste** väärtuste järgi **õpetusandmetes**. Sellest tulenevalt on **Constud** andmebaasis kesksel kohal funktsioontunnuste tabel **FTUNNUSED** ja argumenttunnuste **atribuutide** tabel **ATUNNUSED**. Tabel **FTUNNUSED** on **Constud** andmebaasis esmane objekt — iga uue ülesande ja iga õppevooru alguses loeb **Constud** kõigepealt tabelit **FTUNNUSED**. **Constud** vaatab väljalt [*arvutada*], milliste funktsioontunnustega on tarvis tegeleda, loeb nende kirjete FID numbrid, **andmebaasiobjekti** nime, kus on õpetusandmed ja **näidised** (väli *andmepäring*) ning õppe parameetrid.



Joonis 1-1. Seletavate tunnuste parameetrid tabelis **ATUNNUSED** ning seosed tabelite **FTUNNUSED** ja **ATUNNUSED** vahel. Tabeli **FTUNNUSED** kõigi väljade loetelu on lisas 2.1.



Joonis 1-2. Seletavate tunnuste atribuudid ja seosed tabelite *ATUNNUSED*, *AKIHID* ja *INDEKSID* vahel.

Igal kirjel tabelis *FTUNNUSED* peab olema tabelis *ATUNNUSED* vastav väli, mille nime alguses on F, millele järgneb FID number (joonis 1-1). Näiteks, kui *[FTUNNUSED].[FID] = 32*, siis vastav väli on *[ATUNNUSED].[F32]*. Iga seletava tunnuse puhul peab andmekiht olema määratletud, sest ruumilise muutuja analüüsil loetakse seletavate tunnuste kehtivuse ajalised piirid andmekihtide tabelist. Tunnust kasutatakse kahe ruumikoordinaatidega seotud **vaatluse** vahelise sarnasuse hindamisel vaid juhul, kui tunnuse ajalise kehtivuse vahemik on ülekattes mõlema vaatluse ajalise kehtivuse vahemikuga (joonis 1-3).

Mitteruumilise muutuja analüüsil saadakse aga seletava tunnuste tüüp (nominaalne või arvuline) tunnusega seotud andmekihi tüübist. Andmekihtide teisi atribuute kasutatakse vaid ruumimustri indeksite arvutamisel.

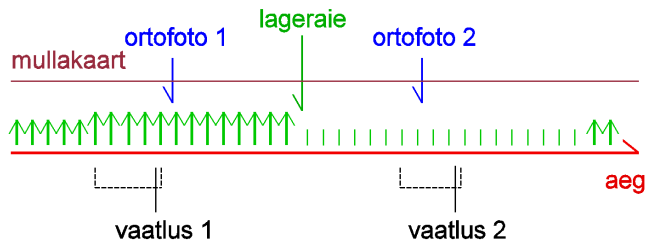
Seletav tunnus peab olema seotud andmekihiga

Teiseks on iga ruumilise tunnuse ehk kohatunnuse puhul kohustuslik indeksi määratlemine, sest indeksite tabelist loetakse, kas ruumilise tunnuse tüüp on nominaalne või numbriline. Mitteruumiliste tunnuste puhul loetakse tunnuse tüüp andmekihi tüübist. Indeksi ID number on kohustuslik isegi valmis arvatud tunnuste kasutamisel, sest tunnuse tüübist sõltub väärtuste kasutamine sarnasuse arvutamisel — klassikoodide erinevuse määramisel ei ole mõtet. Tunnuste teisi atribuute kasutatakse vaid ruumimustri indeksite arvutamisel.

Kohatunnus peab olema seotud ruumimustri indeksiga

Ruumiliste andmete töötlemisel on tunnuse kolmas põhikomponent arvutamise raadius.

Kohatunnuse põhikomponendid süsteemis Constud on:
andmekiht, indeks ja raadius



Joonis 1-3. Kaks vaatlust ajateljel, nende kehtivuse ulatus (katkendjoonega klamber), mullakaardi andmekihi piiramatult kehtivus ning kaks ortofoto andmekiht, mis kehtivad pildistamise kuupäeval. Esimese vaatluse kohta kirjeldamiseks ei kasutata ortofotodelt arvutatud tunnuseid, sest kummagi ortofoto kehtivus ei kattu esimese vaatluse kehtivusulatusega, teise vaatluse kohta kirjeldamiseks saab kasutada mullakaardilt ja teiselt ortofotolt arvutatud tunnuseid.

Andmekihtide tabelist loeb Constud iga tunnusega seotud andmekihi kataloogi nime, tüübi (nominaalne või numbriline), piksli külje pikkuse, baitide arvu piksli kohta, andmekihi eiritava piksliväärtuse ja andmekihi ajalise kehtivuse piirid. Kuna kataloogi nimele liidetakse Constud põhiaknas sisestatud eesliide, siis võib andmekihtide kataloogide ühisosa tabelisse kirjutamata jätta ja sisestada see eesliitena (joonis 1-4). Näiteks kui kõik andmekihid on kataloogi *C:/Constud/Kihid/* allkataloogides, siis tuleb kataloogis *C:/Constud/Kihid/muld/* oleva andmekihi kataloogiks kirjutada lihtsalt *muld*.

Kataloogi eesliide

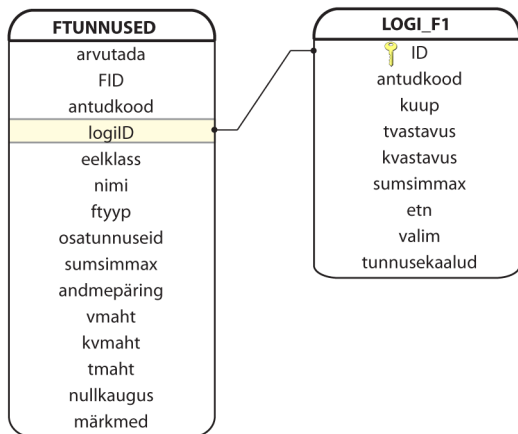
local

FTP login psw

Joonis 1-4. Kataloogi eesliite sisestus Constud põhiaknas.

Logitabelid iga funktsioontunnuse jaoks on tingimata vajalikud nende muutujate jaoks, mida kasutatakse kaalude sobitamisel ja hinnangute arvutamisel. Logitabelitesse salvestatakse iga õppevooru tulemused: etteantud üksus (*antudkood*), kui õpitakse *nominaalse* muutuja igat üksust eraldi, tulemuste salvestamise aeg (*kuup*), *õppevastavus* (*tvastavus*), *kontrollvastavus* (*kvastavus*), otsuse langetamiseks otsitav sarnasuse hulk (*sumsimmax*), mis reguleerib piisavalt sarnaseks loetavate *näidiste* hulka, näidiseks valitud vaatluste arv (*etn*), *valimi maht*

(*valim*), tunnuste optimeeritud kaalud (*tunnusekaalud*). Seni kontrollvastavuse järgi parima õppevooru ID number salvestatakse väljale *[FTUNNUSED].[logiID]*. Väli *ID* seob tabelleid *FTUNNUSED* ja *LOGI_F...* (joonis 1-5).



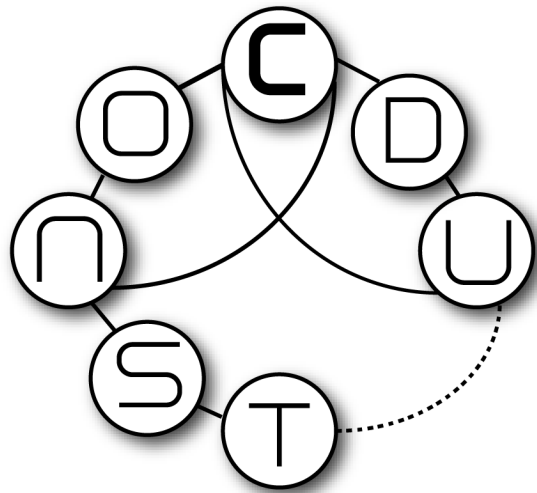
Joonis 1-5. Funktsioontunnuste atribuudid tabelis *FTUNNUSED* ja funktsioontunnuse F1 seos logitabeliga *LOGI_F1*.

Kasutaja võib muuta väärtusi tarkvaraga **Constud** seotud andmebaasitabelites samal ajal kui **Constud** selle sama andmebaasiga suhtleb. Sel viisil saab **tehisõpet** oma tahtmist mööda suunata seda katkestamata.

 **KÜSIMUSED**

1. Kas Constud suudab arvutada kaardimustrit kirjeldavaid indekseid Eesti põhikaardi vektorkujul failidest?
2. Kas Constud abil saab tekitada kaardi, mis näitab kaugust lähimast vooluveest?

ANDMESTRAKTUUR 2



2.1. ANDMEBAAS

Constud toimib vaid koos andmebaasiga, mille struktuur on rangelt ette määratud. Andmebaas võib olla ühes järgmistest vormingutest:

- 1) Access 2000, 2002–2003, 2007–2010,
- 2) MS SQL Server,
- 3) MS SQL Server Express (vabalt allalaaditav:
<http://www.microsoft.com/express/Database/>),
- 4) MS SQL Server Compact Edition.

SQL Server Compact Edition (SQL Server CE) versioon 3.5 SP2 on saadaval aadressil <http://www.microsoft.com/sqlserver/2005/en/us/compact-downloads.aspx>. Uuemat — 4.0 versiooni Constud veel ei toeta, sest versiooni 4.0 on senini arendatud eelkõige veebilahenduste jaoks ja see ei ole veel ladusalt integreeritud SQL Server Management Studio 2008 R2 ja .NET Framework 4.0 keskkonda. SQL Server CE on minimaalse ressursinõudlusega ja kiire andmebaasisüsteem, millel ei ole oma kasutajaliidest. Minimaalsed vahendid pakub SQL Server Management Studio, mida võib täiendada. Vaata näiteks <http://exportsqlce.codeplex.com>. Ülevaade SQL Server CE lisavahenditest on veebilehel <http://erikej.blogspot.com/2009/04/sql-compact-3rd-party-tools.html>. SQL Server Compact 3.5 toetab vaid kuni 4 GB suurust andmebaasifaili ning oluline puudus on, et SQL Server CE ei toeta salvestatud päringuid (andmebaasivaateid). Constud eeldab, et Access või SQL Server Compact andmebaasi fail ei ole parooliga kaitstud.

Access või SQL Server Compact andmebaas ei tohi olla
parooliga kaitstud

Väga suurte andmestike töötlemiseks, ühe ja sama andmebaasi samaaegselt kasutamiseks koos mitme Constud protsessiga ja andmebaasi kasutamiseks üle arvutivõrgu sobib kõige paremini SQL Server. Juhul, kui esmatähtis on töö kiirus, tasuks eelistada SQL Server Compact andmebaasi. Kui on võimalik kasutada vaid tasuta andmebaasitarkvara, siis sobivad SQL Server Express ja SQL Server CE. Constud süsteemiga tutvumisel on lihtsam alustada Access andmebaasiga. Seetõttu lähtuvad ülesannete juhised eelkõige Accessi eripärast.

1. Loo uus tühi andmebaas.



Võimalikud andmebaasivormingud on loetletud eelmises alajaotuses. Algajal on oma arvutis lihtsam töötada Accessi andmebaasiga, andmeserveri tööga tutvumine peaks edasijõudnule huvitavam olema. Otsusta ise, millist vormingut eelistad.

TÜ sisevõrgus saab kasutada Constud vaikimisi ühendust arvutis Vinson oleva

SQL Server ekspressversiooni andmebaasiga *CS1*. Selle andmebaasiga või mõne muu SQL Server tüüpi andmebaasiga töötamiseks väljaspool Constud tarkvara tuleks kasutajal installeerida SQL Server Management Studio (vabalt allalaaditav: <http://www.microsoft.com/express/Database/InstallOptions.aspx> valik *Management Tools*).

Mingisse olemasolevasse andmeserverisse kasutajakonto saamiseks pöördu serveri administraatori poole.

Microsoft Access andmebaasi kasutamise eel võiks kontrollida, kas Accessi teateriba on nähtav ja ega kirjade muutmise ole ära keelatud. Accessi kasutamisel tekkinud tõrgete puhul uuri andmebaasi ja Accessi suvandeid (mis sõltuvad tarkvara versioonist) ja eemalda liigsed piirangud.

2.2. ANDMETABELID

Kui andmebaas on olemas ja Constud kasutajal on sellele administraatori õigustega ligipääs, siis saab hakata andmebaasiobjekte (tabeleid ja vaateid) looma.

Andmebaasivaade on salvestatud päring

Vajalikud tabelid (ja päringud) saab tekitada kas mõnest nädisandmebaasist koopiati tehes või spetsiaalses dialoogiaknas, mida kirjeldatakse järgmises alapeatükis. Nädisandmebaaside juurde on viit Constud veebilehel.

2.2.1. ANDMETABELITE LOOMINE CONSTUD ABIL

Enne tööle hakkamist kontrolli, et sul ei oleks arvutis Constud tarkvara vanemat versiooni kui Constud 2. Kui on installitud varasem versioon, siis kõigepealt eemalda see. Seejärel kasuta veebilehel http://digiarhiiv.ut.ee/Constud2/DW_est.aspx olevaid nuppe programmi käivitamiseks üle võrgu või oma arvutisse installaerimiseks.

Kontrolli, mida kasutab arvuti kümnendkoha eraldajana sinu arvutis. Kui kasutaja arvutis on kümnendkoha eraldajaks koma, siis annab Constud veateate ja palub kümnendkoha eraldaja punktiks vahetada, sest muidu tekib andmebaasiga suhtlemisel tõrkeid. Kümnendkoha eraldajat saab vahetada valides: *Start => Control Panel => Regional and Language Options => Customize => Decimal Symbol*. Muud regionaalseadistused võivad samaks jääda.

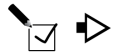
 2. Käivita Constud. Vali tervitusakna vasakust ülanurgast eestikeelne kasutajaliides (joonis 2-1).

Keele valiku järel avaneb Constud põhiaken, millel on alajaotused: *Indeksite arvutus*, *Tehisõpe*, *Teadmiste kontroll*, *Hinnangute arvutus*. Põhiakna alumises osas on andmebaasiühenduse atribuudid (joonis 2-2).

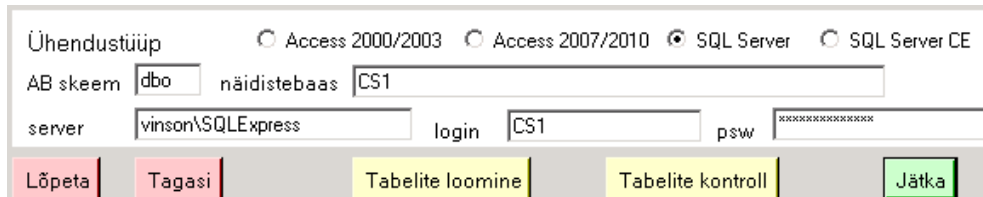


Joonis 2-1. Constud tervitusaken.

3. Vali näidistebaasiks vastloodud andmebaas ja selle tüüp.



Andmebaasi tüübist sõltub ühendustarkvara valik ja seetõttu peab ümarnupust valitud andmebaasitüüp vastama valitud näidistebaasi vormingule.



Joonis 2-2. Andmebaasiühenduse tüübi valik Constud põhiakna eestikeelse versiooni allosas.

4. Vajuta nuppu: Tabelite loomine.



Avaneb tabelite loomise vorm (joonis 2-3). Kui tabelite vormi avanemisel ilmub teade: *Ei saa andmebaasi lugeda*, siis kontrolli, kas andmebaasi nimi ja ühendustüüp said õigesti sisestatud ja kas andmebaasi loomine on ikka lõpule viidud. Tõrke korral võib proovida ka andmebaasiühenduse sulgemist, et samaaegne kasutus või mõni lõpetamata toiming ei takistaks tabelite loomist. Kui tabeli loomise nupp muutus valgeks, aga andmebaasi ei paistnud midagi lisanduvat, siis tihenda ja paranda (*Compact and Repair*) andmebaas.

Tabelite loomise järel tihendada ja parandada andmebaas

A)

Andmebaasi tabelite loomine. NB! Kasutajal peab olema õigus tabelleid luua.

Põhitabelid korraga Tabelid üksikvaaval

Näita veateateid

FTUNNUSED Prognostavate tunnuste loetelu

AKIHID Andmekihide loetelu Kuldsed nupud tähistavad kohustuslikke tabelleid. Tabeli loomise järel muutub nupp valgeks.

ATUNNUSED Seletavate tunnuste loetelu

INDEKSID Ruumimustri indeksite loetelu

PKZõit Kaardilehtede loetelu. Kohustuslik mustriindeksite ja kaartide arvutamisel.

Omasarnasustabelid Tabelele loetelu, kus hoitakse nominaalse tunnuse kategooriate omavahelisi sarnasusi.

Lisa tabelisse FTUNNUSED rida iga prognoositava tunnuse jaoks ja tabelisse ATUNNUSED rida iga kirjeldava tunnuse jaoks

STDEV Standardhälbed. Tabel on kohustuslik, kui standardhälbed kasutatakse.

Vaatluste tabelid **Tunnuste väljad** Vaatluste tabelite nimed loetakse tabeli FTUNNUSED väljalt [andmepäring]

Logitabelid kõigis õpetusandmetes muus tabelis Väli F ja AF tabelis ATUNNUSED peab vastama väljale FID tabelis FTUNNUSED

Väljad F... ja AF... Iga funktsioonitunnuse õppe tulemused salvestatakse eraldi logitabelisse. Parameeter F logitabeli nimes vastab väljale FID tabelis FTUNNUSED.

Sarnasus antud klassiga

Prognoositud väärtused

Vaid juhul, kui omasarnasusi kasutatakse.
Sisesta omasarnasustabeli nimi iga nominaalse argumenttunnuse jaoks, mille puhul omasarnasusi kasutatakse.

Omasarnasustabelid Vajuta nuppu, et luua omasarnasustabelid. **Sulge**

B)

Andmebaasi tabelite loomine. NB! Kasutajal peab olema õigus tabelleid luua.

Põhitabelid korraga Tabelid üksikvaaval

Kõik põhitabelid Näita veateateid

Lisa tabelisse FTUNNUSED rida iga prognoositava tunnuse jaoks ja tabelisse ATUNNUSED rida iga kirjeldava tunnuse jaoks

Standardhälvete tabel on kohustuslik, kui seda kasutatakse.

Siis vajuta nuppu, et luua:

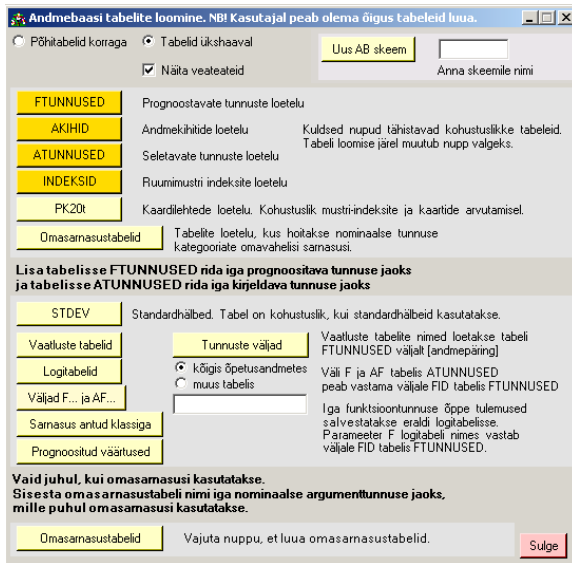
- 1) vaatluste tabelid, Vaatluste tabelite nimed loetakse tabeli FTUNNUSED väljalt [andmepäring]
- 2) iga katse logitabel, Väli F ja AF tabelis ATUNNUSED peab vastama väljale FID tabelis FTUNNUSED
- 3) argumenttunnuste standardhälvete tabel, Iga funktsioonitunnuse õppe tulemused salvestatakse eraldi logitabelisse.
- 4) väljad F ja AF tabelisse ATUNNUSED, Parameeter F logitabeli nimes vastab väljale FID tabelis FTUNNUSED.
- 5) tabel etteantud klassiga sarnasuste jaoks,
- 6) tabel prognoositud väärtuste jaoks.

Loo

Vaid juhul, kui omasarnasusi kasutatakse.
Sisesta omasarnasustabeli nimi iga nominaalse argumenttunnuse jaoks, mille puhul omasarnasusi kasutatakse.

Omasarnasustabelid Vajuta nuppu, et luua omasarnasustabelid. **Sulge**

C)



Joonis 2-3. Tabelite loomise vorm. **A** — Access põhitabelite ükshaaval loomine, **B** — Access põhitabelite korruga loomine, **C** — SQL Server põhitabelite ükshaaval loomine.

SQL Serveri kasutamisel tuleb otsustada, kas lood selle töö jaoks uue andmebaasiskeemi (*schema*) või kasutad olemasolevat. Vaikimisi loodud skeemi nimi on *dbo*. Andmebaasiskeem on andmebaasi struktuurne alljaotus ehk sektsioon. Sama andmebaasi erinevates skeemides võib olla samanimelisi tabeleid, sama andmebaasi eri skeemides võivad kasutajaõigused olla erinevad.

Põhilisi tabeleid saab luua kas korruga või ükshaaval

5. Loo põhitabelid kas ükshaaval iga tabeli nuppu vajutades (joonis 2-3A,C) või korruga nupust Kõik põhitabelid (joonis 2-3B).

Põhitabelite koos või eraldi loomise valimise ümarnupud on akna ülaservas. Põhitabelid on Constud toimimiseks tingimata vajalikud, teisi tabeleid läheb tarvis üksikute tööülesannete juures või teatud tüüpi andmete puhul. Üksikuid tabeleid saab andmebaasi lisada igal ajal.

Kui Constud annab teate, et ta ei saa tabelit luua, siis on enamasti samanimeline tabel andmebaasis juba olemas. Kui tabeli loomine õnnestus, muutub nupp valgeks.

Jäta Constud avatuks. Ava andmebaas ning Access andmebaasi kasutamisel vajuta nuppu *Tihenda ja paranda andmebaas* (*Compact and Repair Database*); SQL Server

andmebaasi puhul vali parema hiireklahviga tabelite osast avanevast rippmenüüst *Refresh*.



6. Sisesta vastloodud tabelisse FTUNNUSED esimese funktsioontunnuse vajalikud omadused.

Esimene harjutus on **ruumimustri** indeksite arvutamine rasterkihist vaatluskohtade ümbruses. Kui kasutad andmebaasitarkvara **Access**, siis kopeeri järgmine SQL lause päringu SQL vaatesse. Päringu SQL vaade on päringu kujundusvaate (*Query => Design View*) üks esitusviisi.

```
INSERT INTO [FTUNNUSED]
([arvutada],[FID],[antudkood],[logiID],[eelklass],[nimi],[ftyyp],
[osatunnuseid],[sumsimmax],[andmepäring],[vmagt],[kvmagt],[tmagt],
[nullkaugus],[märkmeg])
VALUES (False,1,0,0,0, 'indeksite_arvutamine',3,0,1,'VAATLUSKOHAD1',
100,100,10,0,NULL);
```

SQL Server ja SQL Server Express andmebaasi puhul tuleb tabeli nimele lisada andmebaasi ja skeemi nimi, kui need ei ole andmebaasi vaikimisi väärtusteks. Sõnad *True* ja *False* peavad SQL Server ja SQL Server CE tüüpi andmebaasi puhul olema ülakomadega piiratud. Accessi päringutes ei tohi mainitud ülakomasid olla.

SQL Server andmebaasi puhul tuleb tabeli nimele
lisada andmebaasi ja skeemi nimi ning sõnad True ja
False peavad olema ülakomadega piiratud

SQL Server päringu kuju oleks siis näiteks järgmine (andmebaasi ja skeemi nimi täpsusta ise).

```
INSERT INTO [CS1].[dbo].[FTUNNUSED]
([arvutada],[FID],[antudkood],[logiID],[eelklass],[nimi],[ftyyp],
[osatunnuseid],[sumsimmax],[andmepäring],[vmagt],[kvmagt],[tmagt],
[nullkaugus],[märkmeg])
VALUES
('False',1,0,0,0, 'indeksite arvutamine' ,3,0,1,'VAATLUSKOHAD1',100,
100,10,0,NULL);
```

Siin ja järgnevalt mainitud **andmeväljade** formaat ja tähendus on lisas 2, veidi vananenud kujul ka tarkvarasüsteemi **Pidevstudium** kirjelduses (*Remm ja Linder, 2008*). **Sõltuva muutuja** ehk funktsioontunnuse ehkuuritava muutuja tüübid on tabelis **FTUNNUSED** kodeeritud järgmiselt: 0 — multinomiaalne, 1 — mitmemõõtmeline numbriline, 2 — numbriline, 3 — kaheväärtuseline, 4 — multinomiaalse muutuja ühe üksuse eristamine kõigist ülejäänud üksustest.


Funktsioontunnuse koodid: 0 - multinomiaalne, 1 - mitmemõõtmeline numbriline, 2 - numbriline, 3 - kaheväärtuseline, 4 - multinomiaalse muutuja ühe üksuse eristamine kõigist ülejäänud üksustest

7. Ava vastloodud tabel ATUNNUSED ja sisesta sinna esimese argumenttunnuse karakteristikud.

Võid kasutada alltoodud SQL käsku.

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus)
VALUES (1, 'metsa osa', 1, 1, 50, 0, 1, False, 64, False);
```

Tunnuse atribuudi *antudkood* väärtus 64 tähistab siin Eesti põhikaardi põhiala mets. Seega arvutatakse tunnus number 1 (*AID* = 1) nimega metsa osa andmekihist 1 indeksina 1 (antud koodi osa) kasutades kõiki piksliväärtusi (*valim* = 1) 50 ühiku raadiuses (*radius* = 50, *sradius* = 0). Antud koodiks on 64, *eraldisi* ja blokkvalimit ei kasutata.


8. Muuda eelmises ülesandes olevat päringut ja lisa veel kaks seletavat tunnust:  tunnus number kaks (*AID* = 2), mis on kihist number kaks (*KID* = 2) arvutatav keskmine (*indeksID* = 102) kaugusvahemikus 200–300 ühikut vaatluskohast (*radius* = 300, *sradius* = 200) ning lisamaks tunnus number kolm (*AID* = 3), mis on kihist number kolm (*KID* = 3) arvutatav fookuses oleva väärtuse ja 50 ühiku raadiuses oleva ümbruse keskväärtuse vahe (*indeksID* = 115) (*radius* = 50). Lisaks mainitud väljadele tuleks teisele ja kolmandale tunnusele anda mingi nende tunnuste tähendusega seostuv nimi, näiteks *SRTM kõrgus ümbruses* ja *Suhteline kõrgus*.

Tabelite loomise moodul ei nõua andmekihtide kohest määratlemist, aga kuna andmekihid 1 (uusim põhikaart), 2 (*SRTM* kõrgusmudel meetrites) ja 3 (maapinna kõrgus detsimeetrites kahebaidises rastris) olid kahes eelmises päringus mainitud, siis tuleks need järgmiste päringutega lisada andmekihtide tabelisse *AKIHID*. Pane tähele, et kiht 3 on kahebaidises vormingus, sest maapinna kõrguse salvestamisel detsimeetrites saab ühebaidise vorminguga (täisarvud 0 kuni 255) kujutada vaid kõrgust kuni 25,5 meetrini.

SQL käsud on järgmised:

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus)
VALUES (2, 'SRTM kõrgus ümbruses', 2, 102, 300, 200, 1, False, 0, False);
```

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus)
VALUES (3, 'Suhteline kõrgus', 3, 115, 50, 0, 1, False, 0, False);
```

-  **9. Lisa tunnuste 1, 2 ja 3 andmekihid tabelisse AKIHID.** Kataloogid: *PK-uusim*, *SRTM_elev_byte*, *elev_dm* asuvad Ökoloogia ja Maateaduste Instituudi digiarhiivis (<http://digiarhiiv.ut.ee>). Constud suudab digiarhiivi andmekihte lugeda Tartu Ülikooli sisevõrgus ja VPN ühenduse kaudu. Andmekihid võivad asuda ka mujal, kuid peavad olema kindlal kujul korrastatud (*Remm ja Linder, 2008*). Pane viimastest päringutest tähele, et andmekihis *elev_dm* on iga piksli kohta kaks baiti, teistes kihtides üks bait.


Saad kasutada alltoodud päringuid. Andmebaasiväljade vormingud ja tähendused on lisas 2.5.

```
INSERT INTO [AKIHID]
([KID], [pkylg], [kataloog], [nominal], [eira], [alates], [kuni],
[baite], [märkmmed]) VALUES (1, 5, 'PK-uusim', True, 0, null, null, 1, null);
```

```
INSERT INTO [AKIHID]
([KID], [pkylg], [kataloog], [nominal], [eira], [alates], [kuni],
[baite], [märkmmed])
VALUES (2, 100, 'SRTM_elev_byte', False, 0, null, null, 1, null);
```

```
INSERT INTO [AKIHID]
([KID], [pkylg], [kataloog], [nominal], [eira], [alates], [kuni],
[baite], [märkmmed])
VALUES (3, 10, 'elev_dm', False, 0, null, null, 2, null);
```

Kui tunnused ja andmekihid on lisatud, saab jätkata tabelite moodustamist. Enne jätkamist tuleks kõik tabelid sulgeda, et andmebaasiobjektide samaaegne kasutus ei takistaks tabelite loomist ja muutmist.

-  **10. Kui tekitad tabeleid korraga, siis vajuta nuppu Loo, kui ükshaaval, siis nuppe: Vaatluste tabelid, Logitabelid, Väljad F ja AF, Sarnasus antud klassiga, Prognoositud väärtused.**

Tabel etteantud klassiga sarnasuse hoidmiseks luuakse vaid juhul, kui uuritava muutuja tüüp on 4. Seda tüüpi näidisülesanne on peatükis 4.5. Tabeli loomise järel muutub vajutatud nupp heledaks ja tabel ilmub sellesse andmebaasi, millega Constud on ühendatud. Mõne andmebaasitarkvara versiooni puhul tuleb uute objektide nägemiseks vajutada värskendamise nuppu.

Lisada saab veel omasarnasustabelid, milles hoitakse nominaalse tunnuse üksustevahelisi sarnasusi. Kui üksustevahelist sarnasust ei ole määratletud, loetakse see nulliks. Seega võivad omasarnasused ka puududa. Omasarnasustabeli näidis on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *OmasarnasusPK*. Selles tabelis on sarnasused (protsentides) põhikaardi põhialade vahel. Näiteks

põhjala mets (kood 64) ja noor mets (kood 67) vaheliseks sarnasuseks on märgitud 10%. Metsa sarnasus teiste põhjaladega on määratlemata, seega = 0. Klassidevahelised sarnasused võivad olla mingitest reeglitest tuletatud või ka täiesti subjektiivsed. Constud tarkvara eeldab, et sarnasuste väljal on täisarvud vahemikus null ja sada ning arvude tähendust või põhjendatust ei küsi.

Kui kõik õnnestus, siis on tabelite struktuur nüüd andmebaasis olemas. Järgmise tööna tuleks struktuur andmetega täita.

11. Lisa vaatluskohtade koordinaadid vaatluste tabelisse.



Komplekt vaatluskohti on näidisandmete andmebaasis *Constud_6ppe_andmed.mdb*. Kasutagem esmalt soo-neiuvaiba (*Epipactis palustris*) esinemise ja puudumise kohti tabelist *E_palustris*. See tabel tuleks importida või kopeerida Constud andmebaasi. SQL Serverist leiab importfunktsiooni, kui vajutada andmebaasi nime peal paremat hiireklahvi ja valida hüpikmenüüst *Tasks => Import Data*. Access 2007 kasutajaliideses on importfunktsioon menüü rubriigis *Välisandmed*, Access 2003 kasutamisel tuleks menüüst valida *Fail => Välisandmed => Import*. Seejärel avaneb faili valiku aken. Määra faili tüübiks Access 2003, leia fail *Constud_6ppe_andmed.mdb* ja vali imporditavaks objektiks tabel *E_palustris*. Vali import uude tabelisse, **primaarvõtmeks** VID ning imporditud tabeli nimeks *E_palustris*.

SQL Server CE andmebaasi on võimalik importida csv-vormingu kaudu või siis SQL käskude seeria abil, mille saab valmis teha Excelis. Lase csv-vormingusse eksportimisel tunnuste nimed esimesele reale kirjutada.

Vaatluskohtade ülekandmiseks ühest andmebaasitabelist teise on eeskujuks järgmine päring. Praegusel kujul lisab see päring imporditud tabelis *E_palustris* olevad kirjed väljadega VID, *esineb*, X ja Y Constud abil genereeritud tabelisse VAATLUSKOHAD1 väljadele VID, F, X, Y. Välja nimed on kindlam nurksulgudesse kirjutada, sest siis võivad välja nimed sisaldada ka numbreid, sümboleid ja reserveeritud sõnu. SQL Server tüüpi andmebaasi kasutamisel tuleb mõlema tabeli nime ette lisada andmebaasi nimi ja skeemi nimi. Pane tähele, et näidisandmete andmebaasis on esinemine/puudumine väljal *esineb*, Constud andmebaasis on vastab sellele **sõltuva muutuja** väli F.

```
INSERT INTO [VAATLUSKOHAD1] (VID, [F], [X], [Y])
SELECT [VID], [esineb], [X], [Y]
FROM [E_palustris];
```

Värskenda andmebaas ja kontrolli, kas vaatluste atribuudid VID, F, X, Y jõudsid kenasti vaatluskohtade tabelisse. Kui jah, siis võid imporditud tabeli *E_palustris* andmebaasist kustutada. Ära kustuta vaatluskohtade tabelit.

12. Lisa kaardilehtede numbrid ja servakoordinaadid tabelisse PK2ot.



Eesti põhikaardi kaardilehtede numbrid ja servakoordinaadid saab andmebaasi importida nädisandmebaasi *Constud_6ppe_andmed.mdb* tabelist *PK*. Teistsuguse nomenklatuuri kasutamisel tuleks tabelisse *PK20t* kanda sellele teisele nomenklatuurile vastavad kaardilehe numbrid ja servakoordinaadid. Koordinaadid peavad tabelis *PK20t* ja vaatluste tabelis olema samades mõõtühikutes.

Samas *Access* andmebaasis (*SQL Serveri* puhul samas skeemis) ei saa olla kahte samanimelist tabelit, seetõttu on imporditava tabeli nimi *PK*, mitte *PK20t*.

Kaardilehe numbrite ülekandmiseks saad kasutada samasugust *SQL* päringut, nagu vaatluskohtade ülekandmisel — vaid tabelite ja väljade nimed on erinevad. Alltoodud näide sobib kasutamiseks *Accessis*. Kui *SQL Serveris* on mitu andmebaasi või mitu skeemi, siis on kindlam lisada andmebaasi ja skeemi nimi iga tabeli nime ette. *SQL* keele käsud toimivad ilma andmebaasi ja skeemi nimeta vaid juhul, kui andmebaas ja skeem on parasjagu kasutatava ühenduse vaikimisi valikuteks.

```
INSERT INTO [PK20t] ([NR], [min-x], [min-y], [max-x], [max-y])
SELECT [NR], [min-x], [min-y], [max-x], [max-y]
FROM [PK];
```

 **13. Kontrolli, kas kaardilehtede numbrid ja servakoordinaadid jõudsid tabelisse *PK20t* ning kustuta siis tabel *PK*.**

Kontrolli andmebaasi struktuuri ka *Constud* põhiaknas oleva tabelite kontrollimise nupu abil (joonis 2-2). Loodud andmebaasi läheb tarvis selle õpetuse järgmiste sammude läbimiseks.

Tabelite loomise osa kokkuvõtteks võiks järele mõtelda: ehk on edaspidi siiski lihtsam võtta üks olemasolev andmebaas, asendada seal olevad andmed ja kohandada see uue ülesande täitmiseks.

Constud andmebaasi võib kas luua või kopeerida ja kohandada

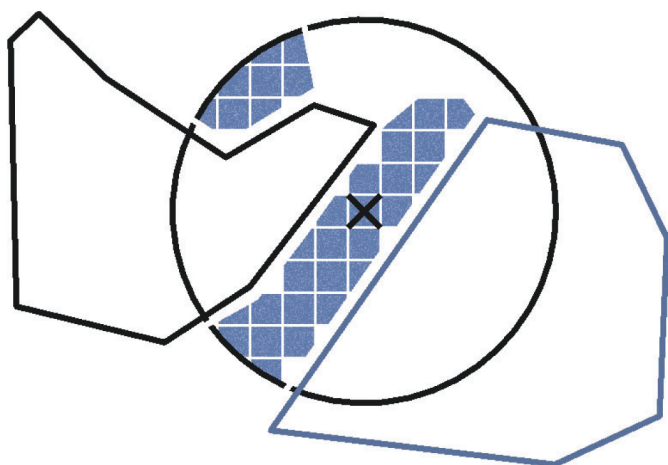


KÜSIMUSED

1. Kas **Constud** suudab toimida koos **Oracle** andmeserveriga?
2. Kas **Constud** suudab toimida koos **MySQL** andmeserveriga?
3. Mitu indeksit on loetletud tabelis *INDEKSID*?
4. Mille poolest erineb nominaalsetest ja numbrilistest andmekihtidest arvutatud indeksite ID number tabelis *INDEKSID*?
5. Milliste tarkvarasüsteemis **Constud** olevate **ruumimustri** indeksite väärtused on nominaalsed? Mis on neil ühist?
6. Millise tabeli millisel väljal on salvestatud **seletavate tunnuste** väärtusi sisaldava **andmebaasiobjekti** nimi?
7. Milline on süsteemis **Constud** kasutatavate omasarnasustabelite maksimaalne arv?
8. Millistel andmebaasiväljadel olevate vahemaade ja koordinaatide mõõtühikud peavad olema omavahel vastavuses **Constud** andmebaasis (loetle tabeli nimi ja välja nimi)?
9. Kuhu salvestatakse seni parima õppevooru ID number lisaks **logitabelile**?
10. Mis ja millistes ühikutes on kirjas väljal *[AKIHID].[pkylg]*?
11. Kas **Constud** võimaldab kasutada andmeid, kus vaatluskohtade koordinaadid on kilomeetrites?
12. Kas **Constud** toetab andmeid, kus vaatluskohtade asukohad on geograafilistes kraadi- ehk polaarkoordinaatides?

MUSTAJINDEKSIO

3



Ruumiandmete kasutamisel Constud süsteemis on argumenttunnused ehk seletavad tunnused andmekihtidest arvatud mustriindeksid. Indeksite loetelu ja arvutuseeskirjad leiad süsteemi kirjeldusest (*Remm ja Linder, 2008*) ning lisast 2.9 ja 3. Mõnede indeksite arvutuseeskiri erineb tavapärasest, et tulemusi saaks salvestada ühebaidilises vormingus. Constud süsteem kasutab senini vaid baidivormingus tunnuseid. Argumenttunnuste väärtused peavad olema täisarvud vahemikus 0–255 isegi siis, kui indeks arvutatakse lühikese täisarvu (väärtusvahemik –32 768 ... +32 767) vormingus andmekihist. Arv 255 tähistab kõigi argumenttunnuste puhul määratlematust — näiteks ei saa määratleda nõlva suunda täiesti horisontaalsel alal.

Kaardina esitatud ruumimustri indeksite näiteid on veebilehel <http://www.geo.ut.ee/LSTATS>. Tarkvara LSTATS arvutab indekseid samal viisil, nagu Constud tarkvaragi.

Seletavate tunnuste väärtused Constud süsteemis on vaid täisarvud vahemikus 0...254, 255 tähistab määramatust või andmete puudumist

14. Muuda välja [FTUNNUSED].[arvutada] väärtuseks True.



Kuna selle juhendi järgi tegutsedes on senini määratletud vaid üks funktsioontunnus, siis seda kirjet tulebki muuta. Kaheväärtuselise muutuja väärtusvarianti True võib andmebaasisüsteem erinevalt kuvada, levinumad kujud on: Yes, 1, –1, Jah, On või linnuke.

Kõik Constud tööülesanded algavad uuritava muutuja määratlemisega tabelis FTUNNUSED. Mustriindeksite arvutamisel on tabelis FTUNNUSED olulised vaid väljad FID, arvutada ja andmepäring. Andmebaasiväli [FTUNNUSED].[arvutada] määrab, milliste uuritavate muutujatega järgnevalt tegeletakse, väli FID on seotud oma väärtusele vastava väljaga tabelis ATUNNUSED (joonis 1-1), väljal andmepäring on selle tabeli või päringu või andmebaasivaate nimi, kust loetakse koordinaadid ja kuhu salvestatakse arvatud tunnused.

15. Ava tabel ATUNNUSED ja märgi kõigi kolme kirje puhul välja F1 väärtuseks True.



Väli on F1 sest selle peatüki ülesannetes on funktsioontunnuseks [FTUNNUSED].[FID] = 1. Funktsioontunnus 1 on nüüd seotud eelnevalt määratletud argumenttunnustega. Kui sa ei leia tabelist ATUNNUSED välja F1, siis järelkult on juhendi samm 10 edukalt läbimata. Kasuta nuppu Väljad F... ja AE... tabelite loomise dialoogiaknas.

3.1. INDEKSID ANDMETABELISSE

Enamik tõrkeid süsteemi Constud kasutamisel tuleneb vigadest andmebaasis. Sellepärast kontrolli iga kord enne indekse arutamist:

- 1) kas funktsioontunnuste tabelis on vajalik rida ja vaid see rida sisse lülitatud;
- 2) kas argumenttunnuste tabeli veerus *F...*, mille numbriline osa vastab sisse lülitatud funktsioontunnusele, on arvutatavad tunnused sisse lülitatud;
- 3) kas väljundtabelis (vaikimisi see andmebaasiobjekt, mille nimi on valitud funktsioontunnuse kirjes tabelis *FTUNNUSED*) on olemas arvutatavate tunnuste veerud (veeru nimi vastab arvutatava tunnuse AID numbrile);
- 4) kas arvutatavate argumenttunnustega seotud andmekihtide andmed ja kataloogi nimi on õiged?

Kontrolli andmebaasi struktuuri ka Constud põhiaknas oleva tabelite kontrollimise nupu abil (joonis 2-2).

- ➔ 16. Käivita Constud. Vali Constud põhiaknast Indeksite arvutuse paneel ja andmekihtide kataloogi eesliide FTP serveris või lokaalvõrgus (joonis 3-1). Vali eelmises peatükis moodustatud Constud andmebaas ja ühendusviis ning vajuta nuppu *Jätka*.

Tartu Ülikooli sisevõrgus ja VPN tunneli kaudu saab kasutada vaikimisi pakutavat FTP ühendust **ÖMI digiarhiivi** FTP osasse. Kasutajad, kellel ei ole ligipääsu Tartu Ülikooli sisevõrku, peavad andmekihid ise ette valmistama. Andmekihtide ettevalmistamise juhend on tarkvarasüsteemi **Pidevstudium** tehnilises kirjelduses (*Remm ja Linder, 2008*).

Indeksite arvutus	Kataloogi eesliide	
<input checked="" type="checkbox"/> Vaatluste tabelisse	<input type="radio"/> local	<input type="text" value="ftp://193.40.8.6/CSkihid/"/>
	<input checked="" type="radio"/> FTP	login <input type="text" value="*****"/> psw <input type="text" value="*****"/>

Joonis 3-1. Valikud Constud põhiaknas ruumilise mustri indekse arutamiseks andmebaasi tabelisse andmekihtide failidest FTP serveris. Sisselogimise nime (*login*) ja parooli (*psw*) lahtrites on varjatud kujul ÖMI digiarhiivi autentimisandmed. Teiste FTP serverite kasutamisel tuleb autentimistunnused ise sisestada.

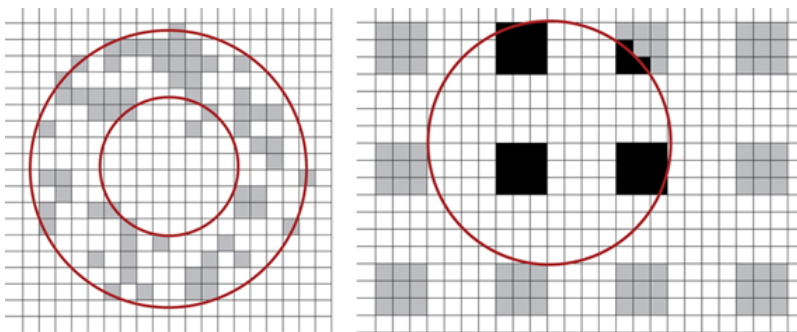
Kontrolli, kas tunnustele 1...3 vastavad väljad [1], [2], [3] vaatluste tabelis täitusid indekse väärtustega kõigil vaatlustel (tabeli kõigis ridades). Meenutagem, et tunnus 1 oli põhikaardi metsa protsentuaalne osa 50 m raadiuses, tunnus 2 oli pinna keskmine kõrgus globaalse **SRTM** mudeli järgi kaugustsoonis 200–300 m, tunnus 3

oli koha suhteline kõrgus 50 m raadiuses oleva ümbruse keskmise suhtes pluss 100 (positiivsete pinnavormide puhul on tunnuse 3 väärtused >100, negatiivsete puhul <100). Kahe viimase tunnuse mõõtühikud on samad, mis vastaval tunnusel, SRTM andmekihis on kõrgused meetrites, andmekihis *elev_dm* detsimeetrites.

3.1.1. VALIMI KASUTAMINE

Keeruka arvutuseeskirjaga indekseid arvutamine suure hulga piksliväärtuste järgi paljudes kohtades võib olla üsna aeganõudev. Arvutuste maht on eriti suur nende indekseid puhul, mis nõuavad väärtuste võrdlemist pikslipaarides, sest N objektist saab moodustada $N(N-1)$ paari. Õnneks piisab enamikul juhtudel mõõduka suurusega valimist, et saavutada vajalik esinduslikkus ja jätta juhuslik viga vastuvõetavatesse piiridesse. Sel juhul ei ole mõistlik liigselt aega kulutada arvutamaks indekseid väärtust tuhandetest või veelgi suuremast hulgast pikslitest iga vaatluskoha ümber. Constud võimaldab kasutada juhuvalimit ja blokkvalimit, mõlemat valimit saab kombineerida ümar ja sõõrik-kerneliga (joonis 3-2).

Juhuvalimi oodatav osakaal määratakse neljabaidise murdarvuna Constud andmebaasi väljal *[ATUNNUSED].[valim]*, blokkvalimi kasutus märgitakse väljale *[ATUNNUSED].[blokkvalim]*. Blokkvaliku puhul valitakse 3×3 piksli suuruseid blokke, mis tagab suurema hulga külgnevaid piksleid kui sama osakaaluga juhuvalik. Blokkide vahe sõltub väljal *valim* määratud valimi osast sellisel viisil, et blokkidesse valitud pikslite osa vastaks valimi osale nii täpselt kui võimalik. Blokkvalimit tuleks kasutada arvutades indekseid, mis nõuavad naaberpikslite võrdlemist olukorras, kus valimi osa peaks olema suhteliselt väike, sest kernelis on palju piksleid. Naaberpikslite võrdlemist nõuab näiteks indeks number 114 (üleminekute sujuvus) ja indeks number 105 (autokorrelatsioon kaheksa naaberpiksli järgi).



Joonis 3-2. Pikslite juhuvalim sõõrik-kernelis ja blokkvalim ümarkernelis.

Uue tunnuse arvutamiseks tuleb Constud andmebaasi lisada mõned väljad ja parameetrid.

17. Lisa tabelisse ATUNNUSED kahe tunnuse atribuudid, mis erinevad tunnusest 1 vaid valimi mahu poolest; lülita nendel sisse väli F..., mille number vastab tabelis



FTUNNUSED reale, millel [arvutada] = True.

Tunnus 1 arvutati põhikaardi põhialade viiemeetrise küljega piksliteks rasteriseeritud kihist 50 m raadiuses vaatluskoha ümbruses. Viiemeetrise piksleid on 50 m raadiusega ümaras kernelis 317, kui kaardi serv osa kernelist ära ei löika. Järgnevate päringutega saab lisada tunnuse 1 modifikatsioonid — ühe puhul on valimi mahuks 50%, teise puhul 10%.

```
INSERT INTO [ATUNNUSED] (AID, nimi, KID, indeksID, radius,
sradius, valim, blokkvalim, antudkood, eraldistekasutus,
eraldistekiht, BV_kataloog, asendab, divisor, F1)
VALUES (4, 'metsa osa 50%',1,1,50,0,0.5,False,64,False,null,null,
null,null,True);
```

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor, F1)
VALUES (5,'metsa osa 10%',1,1,50,0,0.1,False,64,False,null,null,
null,null,True);
```

Kui kasutad **SQL Server** andmebaasi, siis lisa ülakomad sõnade *True* ja *False* ette ja järele ning andmebaasiskeemi nimi tabeli nime ette.

```
INSERT INTO [CS1].[C-Est].[ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor, F1)
VALUES (4,'metsa osa 50%',1,1,50,0,0.5,'False',64,'False',null,null,
null,null,'True');
```

```
INSERT INTO [CS1].[C-Est].[ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor, F1)VALUES (5,'metsa osa 10%',1,1,50,0,0.1,'False',64,False,
null,null,null,null,'True');
```

SQL Server CE puhul kasuta ülakomasid sõnade *True* ja *False* ees ja järel, kuid andmebaasiskeemi ei ole reeglina tarvis (kui sa just ise ei ole Constud andmebaasi mitut skeemi tekitanud).



18. Lülita välja varasemate tunnuste (AID < 4) arvutus [ATUNNUSED].[F1] = False.

Access:

```
UPDATE [ATUNNUSED] SET [F1]=False WHERE [AID]<4;
```

SQL Serveri puhul lisanduvad ülakomad ja skeemi nimi näiteks järgmisel kujul:

```
UPDATE [CS1].[C-Est].[ATUNNUSED] SET [F1]='False' WHERE [AID]<4;
```

19. Lisa vaatluste tabelisse väljad [4] ja [5].



Vaatluste tabel on see tabel või **andmebaasivaade**, mille nimi on väljal [FTUNNUSED]. [andmepäring]. Vaata ja kasuta alltoodud päringuid või kasuta nuppu Tunnuste väljad tabelite loomise aknas.

Access:

```
ALTER TABLE [VAATLUSKOHAD1]
ADD [4] byte NULL, [5] byte NULL;
```

SQL Serveri puhul pane tähele, et baidiformaadi tähiseks on *tinyint*.

```
ALTER TABLE [CS1].[C-Est].[OBSERVATIONS1]
ADD [4] tinyint NULL, [5] tinyint NULL;
```

Kontrolli, kas väljad [4] ja [5] on olemas.

20. Käivita Constud arvutamaks tunnuste 4 ja 5 väärtused.



Võrdle tulemusi tunnuse number 1 väärtustega, mis on saadud valimit kasutamata.

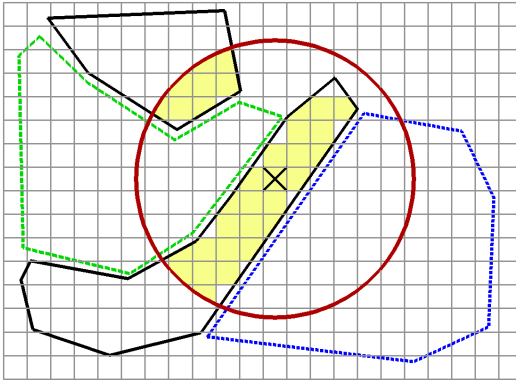
3.1.2. INDEKSITE ARVUTAMINE ERALDISTE PIRES

Mõnikord on põhjust piirata mustriindeksi arvutamise aluseks olevaid piksleid mingite etteantud eraldistega. Näiteks puistu võradelt peegeldunud kiirguse keskmise hulga arvutamisel metsa serva juures oleva vaatluskoha ümbruses tuleks kasutada vaid metsa piires olevaid kaugseirekujutise piksleid. Kui osa kasutatavas raadiuses olevaid kaugseirekujutise piksleid pärineb lagedalt alalt, siis nende kaasamine moonutaks võrade keskmist näitajat. Samalaadne probleem on, kui kohta kirjeldavad tunnused tuleb kaardi- ja kaugseireandmetest arvutada kirjelduse kehtivuse piires. Seega piiravate eraldiste kasutamisel arvatakse valimisse kõik pikslid, mis on etteantud raadiuse piires ja mis kuuluvad eraldiste kihis vaatluskoha piksliga samasse üksusesse (joonis 3-3).

21. Lisa tunnused 6 ja 7 arvutamaks piksliväärtuste standardhälvet Landsat 7 ETM+ kujutisest 23.05.2005 kasutades tunnus 7 puhul piiravate eraldistena põhikaardi põhialasid.



Selle Landsat kujutise failid on ÖMI digiarhiivi kataloogis ETM-186_019-23-05-2005-B8. Mõlema tunnuse puhul peaks piksli külg olema 10 m ja kerneli raadiuseks tuleks määrata 50 m. Tunnus 7 puhul kasutatakse piiravaid eraldisi andmekihist number 1 (ülesanne 9). Standardhälve on Constud indeks nr 103 (tabel INDEKSID). Constud andmebaasi täiendused on sarnased sellele, mida tegid ülesannetes 9 ja 17–19. Vaata hoolega alltoodud näidispäringuid, mis sobivad Accessis kasutamiseks. SQL Server andmebaasi puhul tuleb süntaksit veidi muuta (skeemi nimi, ülakomad sõnade True ja False eel ja järel, tinyint tähistab ühebaidilist vormingut).



Joonis 3-3. Eraldiste piirav mõju pikslite valikule selle tunnuse jaoks etteantud raadiuses (punane ring). Roheline, sinine ja must joon piiravad eraldisi, iga värv tähistab omaette klassi eraldiste kihis. Vaatluskoht (x) on sinise klassi alal. Tunnuse arvutamisel kasutatakse vaid kollasega tähistatud piksleid. Piksleid on valimisse arvatud keskkoha paiknemise järgi.

Kanna Landsat kujutise andmekihi kirje Constud andmebaasi:

```
INSERT INTO [AKIHID]
([KID],[pkylg],[kataloog],[nominal],[eira],[alates],[kuni],[baite],
[märkmed])
VALUES (4, 10,'ETM-186_019-23-05-2005-B8', False, 0, null, null, 1,
null);
```

Lisa tunnuste määratlused:

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor,F1)
VALUES (6, 'SD of ETM+image', 4, 103, 50, 0, 1, False, 0, False, null, null,
null, null, True);
```

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor,F1)
VALUES (7, 'SD of ETM+image', 4, 103, 50, 0, 1, False, 0, True, 1, null, null,
null, True);
```

Lisa väljad arvutatavatele väärtustele:

```
ALTER TABLE [VAATLUSKOHAD1] ADD [6] byte NULL, [7] byte NULL;
```

Kontrolli, kas väljad [6] ja [7] on olemas.

Lülita välja eelmiste tunnuste arvutamine.

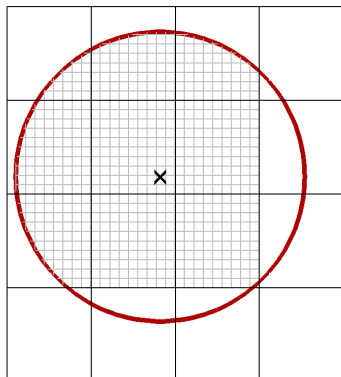
```
UPDATE [ATUNNUSED] SET [F1]=False WHERE [AID]<6;
```

22. Käivita Constud arvutamaks tunnuste number 6 ja 7 väärtusi.

Võrdle arvutatud väärtusi.

3.1.3. ÕMBLUSTETA ANDMEKIHID

Kui Constud indekse arvutamise osast on valitud kaardilehtede liitmine õmblusteta andmekihiks (joonis 3-4), kaasab Constud naaberkaardilehe piksleid, mis jäävad tunnuse jaoks määratletud raadiuse ulatusse.



Joonis 3-4. Pikslite kaasamine esimese astme naaberkaardilehtedelt õmblusteta arvutuse korral. Must rist tähistab vaatluskohta, punane ring tunnuse arvutamise raadiust, mustad ruudud kaardilehti, hall võrgustik arvutusse kaasatud piksleid. Teise astme naaberlehtedelt piksleid ei kaasata.

23. Lisa identsete atribuutidega (va AID) tunnused 8 ja 9. Las arvutatav indeks olla põhikaardi põhialade (kiht 1) arv (indeksID = 6) 100 m raadiuses.



Tunnuste lisamine:

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor,F1)
VALUES (8, 'põhialade arv',1,6,100,0,1,False,0,False,null,null,null,
null,True);
```

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, asendab,
divisor,F1)
VALUES (9, 'põhialade arv',1,6,100,0,1,False,0,False,null,null,null,
null,False);
```

Lisa tunnuste väljad vaatluste tabelisse.

```
ALTER TABLE [VAATLUSKOHAD1] ADD [8] byte NULL, [9] byte NULL;
```

Kontrolli, kas väljad [8] ja [9] on olemas.

näidisandmete andmebaasi Constud_6ppe_andmed.mdb tabelist Prognooositavad_kohad prognoositavate väärtuste tabelisse, mille nimi peaks seda juhendit järgides olema HINNANGUD_1.

Samalaadset ülesannet täitsid näiteks punktide 11 ja 12 juures.

27. Lisa prognoositavate väärtuste tabelisse tunnuste väljad.



Selle juhendi järgi toimides peaks see tabel juba hõlmama kolme esimest tunnust. Ülejäänud saad lisada kasutades nuppu *Tunnuste väljad* tabelite loomise aknas ja kirjutades nupu all olevasse lahtrisse väljundtabeli nime või alltoodud päringuga. Vaikimisi on lahtris *õpetusandmete* tabeli nimi, asenda see prognoositavate väärtuste tabeli nimega (*HINNANGUD_1*).

```
ALTER TABLE [HINNANGUD_1] ADD
[4] byte NULL,
[5] byte NULL,
[6] byte NULL,
[7] byte NULL,
[8] byte NULL,
[9] byte NULL;
```

Kontrolli, kas tunnused 1...9 on tabelis olemas.

28. Lülita sisse kõigi tunnuste arvutamine.



```
UPDATE [ATUNNUSED] SET [F1] = True;
```

29. Käivita Constud, vali paneel Indeksite arvutus ja sellelt valik Prognoosi tabelisse (joonis 3-6) ja sisesta andmebaasiobjekti nimi, mis selle juhendi järgi töötades peaks olema HINNANGUD_1.



Kontrolli arvutuse tulemusi.

Indeksite arvutus	Tehisõpe	Teadmiste kontroll
<input type="checkbox"/> Vaatluste tabelisse		
<input checked="" type="checkbox"/> Prognoosi tabelisse		HINNANGUD_1

Joonis 3-6. Constud valikud indeksite arvutamisel prognoositabelisse.

Constud pakub eraldi valikuna võimalust indeksite arvutamiseks etteantud üksusega sarnasuste tabelisse. See on vaid üks täiendav võimalus sisestada väljundtabeli nime. Andmete ettevalmistuses ja arvutamises eelnenud ülesannetes esitatuga võrreldes vahet ei ole.


3.2. INDEKSID RASTERFAILI

Indeksite rasterfaili arvutamisse võib lisaks sisend- ja väljundfailidele olla kaasatud pikslite valikut piiravate eraldiste andmekiht. Nende kolme andmekihi paiknemise osas kehtivad mõned piirangud. Esiteks — väljundkataloog peab olema kohalik (kas kasutaja arvutis või sinna kaardistatud võrgukataloog, kuhu sul on kirjutamise õigus). Teiseks — ka indeksite arvutamise sisendkataloog peab olema kohalik, sest kõik sisendkataloogis olevad rst-laiendiga failid teisendab Constud automaatselt üksteise järel väljundfailideks. FTP kataloogis on enamasti rohkem faile, kui on tarvis teisendada. Ei ole mõtet teisendada kõiki kaardilehti, kui väljundkaart jääb vaid ühe kaardilehe piirsesse.

Jääb vaid kolm võimalikku kihtide paiknemise kombinatsiooni: 1) piiravad eraldised on lokaalkataloogis, 2) piiravad eraldised on FTP kataloogis, 3) piiravaid eraldisi ei kasutata. Esimesel juhul vali Constud indeksite arvutamise osast *local* ja võid kasutada kataloogi nime eesliidet, mis kehtib nii eraldiste kui ka sisendfailide kataloogi nime puhul. Teisel juhul vali *FTP* (joonis 3-7) ja määra kataloogi nime eesliide, mis kehtib vaid eraldiste andmekihile. Kui valitud on *FTP*, siis indeksi arvutamise lähteandmete andmekihi kataloogi jaoks eesliide ei kehti; lähteandmete kataloogi täisnimi peab olema väljal *[AKIHID].[kataloog]*.

-  30. Loo oma arvutisse uus kataloog sisendfailide jaoks; anna kataloogile nimi, näiteks C:\soil_10.

Constud arvutab indeksid kõigist rst-laiendiga failidest selles kataloogis.

-  31. Kopeeri näidisanemete hulgast failid 5434.rst, 5434.rdc, 5444.rst, 5444.rdc oma sisendfailide kataloogi.

Allalaaditud failides on kaardilehtede 5434 ja 5444 kümne üksuse tasemele üldistatud mullaliigid Eesti 1: 10 000 mullakaardilt. Madalsoomullad on tähistatud koodiga 10.

-  32. Lisa tabelisse AKIHID kirje, mis esindab Eesti 1: 10 000 mullakaarti 10 üldistatud mullaliigina.

Sisesta andmekihi väljale *kataloog* sama nimi, kuhu tõmbasid mullakaardi rst-failid. *[AKIHID].[kataloog]* = {kataloog, mille löid sammuga 30}. Võid kasutada järgmist päringut, aga ära unusta muutmast kataloogi nime C:\soil_10 mullakaardi rst-failide kataloogi täisnimeks su enda arvutis. Kataloogi täisnimi algab salvestusseadet või selle osa tähistava tähe ja kooloniga.

```
INSERT INTO [AKIHID]
([KID],[pkylg],[kataloog],[nominal],[eira],[alates],[kuni],[baite])
VALUES (5,10, 'C:\soil_10', True,0,null,null,1);
```

33. Lülita välja olemasolevate tunnuste arvutamine muutes tabelit ATUNNUSED.



Arvuta rastrisse vaid uus tunnus number 10.

```
UPDATE [ATUNNUSED] SET [F1] = False;
```

34. Lisa uus rida tabelisse ATUNNUSED. Seo see rida üsja lisatud andmekihiga ja määra tunnuse atribuutideks madalloomulla osa 100 m raadiuses sama põhiala (kiht 1) piires, millel on vaatluskoht.

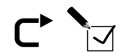


Madalloomulla kood andmekihis 5 on 10. Määra arvutatavate väljundfailide kataloog (väli *BV_kataloog*). Väljundkataloog peaks erinema tabelis *AKIHID* kirjas olevast sisendkataloogist. Kui sisend- ja väljundkataloog on sama, siis failid kas kirjutatakse üle või tekib kirjutamistõrge. Uue rea tabelisse *ATUNNUSED* saad lisada alltoodud päringuga. Ära unusta asendamast näidises olevat kataloogi nime *C:\soil_10\output* väljundkataloogi nimega oma arvutis. Väljundkataloogi puhul peab kirjas olema täisnimi — eesliidet ei saa siin kasutada. Kui tabelis olevat väljundkataloogi arvutis ei ole, siis üritab Constud seda ise luua.

Kui eraldistekiht loetakse FTP serverist, siis peab ka lähtefailide kataloogil täisnimi kirjas olema. Kataloogi nime eesliidet saab indekse rastrisse arvutamisel kasutada vaid siis, kui nii eraldistekiht kui ka lähteandmete kiht on kliendi arvuti kataloogis. Kui eraldistekiht loetakse FTP serverist, peab Constud põhiaknas olema valitud FTP (joonis 3-7). FTP ja kohalike kihtide kasutamisest on juttu peatüki 3 alguses, tunnuste põhilistest komponentidest peatükis 2.1.

```
INSERT INTO [ATUNNUSED]
(AID, nimi, KID, indeksID, radius, sradius, valim, blokkvalim,
antudkood, eraldistekasutus, eraldistekiht, BV_kataloog, F1)
VALUES (10, 'madalloomulla osa', 5, 1, 100, 0, 1, False, 10, True, 1,
'C:\soil_10\output', True);
```

35. Käivita Constud. Vali indekse arvutuse sektsioon ja vali Rasterfaili. Sisesta piksli küljeks 50 m, et arvutatavate pikslite hulk oleks mõõdukas ja arvutus oleks kiire.




Kui piiravate eraldiste kiht asub FTP serveris, siis vali ümarnupp *FTP* ja kasuta selles serveris oleva kataloogi eesliidet (joonis 3-7). Kui aga ka eraldistekiht on kohalikus kataloogis või kui piiravaid eraldisi ei kasutata, siis jäta kataloogi eesliite lahter tühjaks, sest kirjutasid väljale *[ATUNNUSED].[BV_kataloog]* kataloogi täisnime. Valitud tunnused arvutatakse kõigest sisendkataloogis olevatest rst-laiendiga failidest.

3.3. INDEKSID KAHEBAIDISE TÄISARVU VORMINGUS ANDMEKIHIST

Kõigi tunnuste väärtusi hoitakse Constud andmebaasis ühebaidises vormingus (täisarvud vahemikus 0...255). Baidivorming kanali kohta on kaugseire andmetes traditsiooniline ning sellest reeglina piisab ka kategooriliste muutujate kaardistamiseks. Mõnedel juhtudel täisarvudest siiski ei piisa. Näiteks võimaldab ühebaidine vorming salvestada vaid väga robustset kõrgusmodelit.

Maapinna kõrgus detsimeetrites on reeglina piisav täpsus, et kaasata kõrgusandmed maakatte kaardistusse. Detsimeetrites kõrguse salvestamiseks enamikul Maa pinnast sobib kahebaidise täisarvu vorming. Sellises vormingus andmekihti tegelikult juba kasutasime ülesandes 9. Vaata, tabelis AKIHID on kirje $KID = 4$ ja selle väljal $[AKIHID].[baite] = 2$. See tähendab, et piksliväärtused selles kihis on kahebaidises vormingus. Sellest kihist tuli arvutada tunnus number 3 (suhteline kõrgus). Suhtelised kõrgused salvestati ühebaidise tunnusena detsimeetrites, kuna nimetaja (väli *divisor*) abil teisendamist ei tellitud. Teisendamine ei olnud vajalik, kuna maapinna kõrguse erinevused uurimisalal ei ole eriti suured.

Kui aga arvutada samast kahebaidisest andmekihist maapinna keskmist kõrgust mingi raadiusega ümbruses, saaksime valdavalt ühetaolised tulemused 254 (255 on reserveeritud määramatusele ja andmete puudumisele). Võimaldamaks kahebaidisest kihist arvutatava tunnuse automaatset konverteerimist ühebaidisesse vormingusse, on Constud andmebaasis väli $[ATUNNUSED].[divisor]$. *Divisor* ehk jagaja mõjutab tunnuste väärtusi nii tabelisse kui ka rasterfaili arvutamisel. Tunnuse arvutamisel kahebaidisest andmekihist toimub arvutus kahebaidiste arvudega kuid tulemus teisendatakse ühebaidisesse vormingusse. Kui nimetaja väli on tühi, loetakse selle väärtuseks 1. Nimetajat kasutatakse vaid siis, kui andmekihi piksliväärtused on rohkem kui ühebaidised.

37. Kopeeri tabelist ATUNNUSED rida, millel AID = 3, muuda AID väärtuseks 11 ja välja $[ATUNNUSED].[divisor]$ väärtuseks 10. Anna uuele tunnusele nimeks: suhteline kõrgus meetrites. Lülita välja kõigi teiste tunnuste arvutus, telli vaid selle tunnuse arvutus määrates $[ATUNNUSED].[F1] = True$. 

38. Lisa tunnus 11 ka vaatluste tabelisse ja hinnangute tabelisse. 

39. Käivita Constud arvutamaks suhtelisi kõrgusi meetrites vaatluste tabelisse ja prognooside tabelisse. 

Võrdle väärtusi tunnusel number 3 (suhteline kõrgus detsimeetrites) ja tunnusel number 11 (suhteline kõrgus meetrites). Pane tähele, et suhtelisele väärtusele lisatakse konstant 100, et saaks nii positiivseid kui ka negatiivseid tulemusi salvestada baidivormingus (täisarvudena 0...255).



40. Moodusta uus tunnus AID numbriga 12 tabelisse ATUNNUSED. Lisa sellele tunnusele vastav väli tabelisse HINNANGUD_1. Lase Constud arvutab selle tunnuse väärtused.

Meenuta, et ruumilise tunnuse määratluse peamised komponendid on andmekiht, indeks ja raadius. Indeksite loetelu on tabelis *INDEKSID*, mille väli *nomkiht* näitab, kas indeksit saab arvutada nominaalsest andmekihist või numbrilisest, väli *nomind* näitab, kas arvutatud tunnus on sisult **nominaalne** või arvuline. Nagu indeksite tabelist selgub, annavad nominaalse väljundi vaid ühel või teisel moel arvutatud **moodid**.

Numbrilise tunnuse loomiseks saad kasutada Constud andmebaasi juba lisatud andmekihti number neli, mis on juba registreeritud tabelisse *AKIHID* ja mis viitab **ÖMI digiarhiivi** kataloogile *ETM-186_019-23-05-2005-B8*. Võid ka digiarhiivist mingi muu andmekihi valida ja tabelisse *AKIHID* registreerida. Digiarhiivist omavalitud andmekihis kasutatava pikslisuuruse saad arvutada faili suuruse järgi. Kõik digiarhiivi CS-kihtide valdkonna rst-failid on 10 x 10 km kaardilehtedena ja ühebaadises vormingus.

➡ **41. Lase Constud arvutab väärtused uuele tunnusele number 12.**

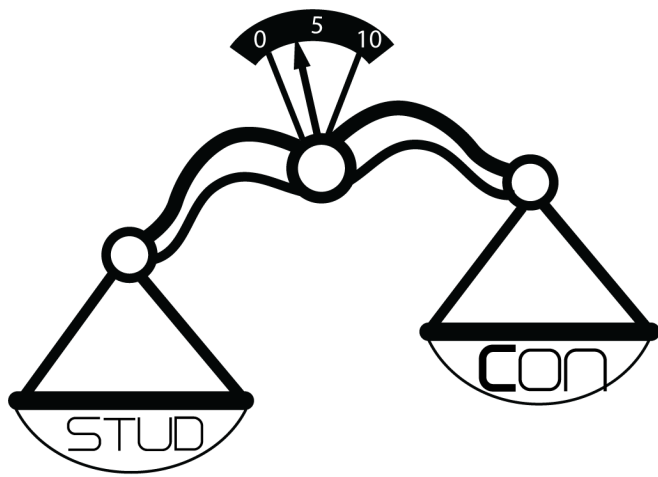


KÜSIMUSED

1. Kas 50 m raadiuses arvutatud Constud indeks 115 järgi on soo-neiuvaiba vaatluskohtade hulgas sagedasemad positiivsed või negatiivsed pinnavormid? Vaatluskohad on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *E_palustris*, tunnused arvutasid tabelisse *VAATLUSKOHAD1*.
2. Tunnused 1, 4 ja 5 kirjeldavad metsa osa 50 m raadiuses. Erinevad need tunnused vaid **kernelis** olevate **pikslite** valimi osakaalu poolest. Koosta ja esita punktdiagramm Excelis, mis kujutaks seost tunnuste 1 (horisontaalteljel) ja 5 (vertikaalteljel) vahel näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *E_palustris* olevates vaatluskohtades. Tunnused arvutasid tabelisse *VAATLUSKOHAD1*.
3. Tunnus 7 erineb tunnusest 6 vaid andmekihi 1 **eraldiste** kasutamise poolest. Mitmes vaatluskohas on tunnuste 6 ja 7 väärtused erinevad? Vaatluskohad on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *E_palustris*, tunnused arvutasid tabelisse *VAATLUSKOHAD1*.
4. Kui suur on eelpool kasutatud Landsat kujutise piksliväärtuste **standardhälbe** (tunnus 6) keskmine kohtades, mille 50 m ümbruses ei ole metsa ja kohtades, mille 50 m ümbruses on vaid mets? Vaatluskohad on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *E_palustris*, tunnused arvutasid tabelisse *VAATLUSKOHAD1*. Metsa osa 50 m raadiuses arvutasid väljale [1].
5. Mitmes kohas on **õmblusteta andmekihist** arvutatud tunnuse 9 väärtus suurem kui tunnusel 8 (sama tunnus arvutatuna igal kaardilehel eraldi)? Vaatluskohad on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *E_palustris*, tunnused arvutasid tabelisse *VAATLUSKOHAD1*.
6. Kaardilehe 5434 keskel on Pühajärv. Ava enda genereeritud kaardileht programmiga *Idrisi* ja vaata, mis on tunnuse number 10 (madalsoomulla osa 100 m raadiuses) väärtused Pühajärve ja teiste järvede ning samuti asulate kohas. Kuidas tõlgendada sellist väärtust arvutunnuse puhul, mis peaks olema vahemikus 0–100%?
7. Esita tunnuse number 12 atribuudid tabelist *ATUNNUSED* ja iseloomusta arvutatud tulemusi. Selle tunnuse koostasid ülesandes nr 40 ja arvutasid ülesandes nr 41.

КАЯЛУО
KAYALUO

4
4



4.1. TEHISÕPE EHK KAALUDE AUTOMATISEERITUD SOBITAMINE PROGRAMMIS CONSTUD

Tarkvaralahenduses **Constud** kasutatakse näidistele tuginevat ehk sarnasusele tuginevat järeldamist, mis väldib üldistusi ja eristub selle poolest mudelitele tuginevast järeldamisest. Sarnasuse järgi järeldamisel otsitakse olemasolevatest **õpetusandmetest** prognoositavale vaatlusele kõige enam sarnanevaid eeskujusid ehk etalone ehk **näidiseid**. Prognoositava muutuja väärtuseks võetakse selle muutuja väärtused näidistel.

Sarnasust mõõdetakse objektide (**vaatluste**, kohtade, juhtumite) vahel ning neid võrreldavaid objekte kirjeldavate tunnuste järgi. Sarnasusele tugineva järeldamise korral on põhiküsimusteks — mille poolest sarnane ja kellega sarnane — ehk kui suurt tähendust omistada ühele või teisele tunnusele ja ühele või teisele näidisele. Kui sarnasuse järgi otsustaja valib ise, milliste tunnuste osas sarnasust ta oluliseks peab, siis on tulemus paratamatult subjektiivne. Tunnuste ja näidiste kaalude automatiseeritud sobitamise eesmärk on sellist subjektiivsust vältida. Kasutaja otsustada on paraku ikkagi, milliseid õpetusandmeid ja **Constud** **suvandeid** kasutada.

Sarnasust mõõdetakse objektide vahel ja tunnuste järgi.
Ühe tunnuse poolest sarnased objektid võivad erineda
teise tunnuse poolest

Tehisõpe tarkvaralahenduses **Constud** on piiramatult korduv järk-järguline selliste kaalude otsimine tunnustele ja näidistele, mille kasutamisel oleksid prognoosid kõige täpsemad. Suure vaatluste ja tunnuste hulga korral on enamik õpetusandmeid hinnangute arvutamiseks mittevajalikud — enamasti piisab vähem kui kümnest tunnusest ja vähem kui paarisajast näidisest. Ülejäänud tunnused ja vaatlused saavad prognoosivas komplektis kaaluks null. Seega, tehisõppe tulemus on kaalude komplekt, milles on kaal igale tunnusele ja kaal igale õpetusvalimis olevale vaatlusele. **Constud** tehisõpe võib pikalt jätkuda, parajasti leitud parim kaalude komplekt on vaid seni parim prognoosiv komplekt. Kuna murdarvuliste kaalude kombinatsioonide hulk on sisuliselt lõpmata suur, siis võib alati leiduda mingi prognoosimiseks veel sobivam variant.

Parim prognoosiv komplekt koosneb tunnuste ja
näidiste kaaludest

Constud tehisõpe koosneb õppevoorudest ehk **õppekordustest**. Ühe õppevooru sees rakendatakse kaalude sobitamise erinevaid mooduseid. **Constud** tehisõppe erinevaid tüüpe täitvad alamprogrammid toimivad kindlas järjekorras (joonis 4-1). Algul valitakse tunnuseid, siis järgneb tunnuste kaalumise, näidiste valik, näidiste

kaalumine, tunnuste ja vaatluste aktuaalsuse muutmine ning õppevoorude tulemuste salvestamine. Kasutaja saab suvanditest valida, mitme korduse jooksul sobitada tunnuste kaale ja mitu korda vaatluste kaale (peatükk 4.2). Constud tehisõppe alamprogrammid kasutavad järgmisi lihtsaid algoritme.

Constud õppevoorud sisaldavad erinevaid kaalude muutmise osi. Iga osa sisaldab korduvat toimingut

Tunnuste valik. Proovitakse tunnuseid ükshaaval järjest väheneva kaaluga lisada. Sarnasust hinnatakse kõigi vaatluste järgi, sest näidiseid ei ole veel valitud. Parim tunnuste kogum (mis enamasti tähendab osade tunnuste kasutusest välja jätmist) jätkab järgmises moodulis.

Tunnuste kaalumine. Eelmises etapis valitud tunnuste kaale sobitatakse neid katseliselt väikese juhusliku väärtuse võrra suurendades ja vähendades. Kui kaalumuutus muudab hinnangud täpsemaks, siis kasutatakse muudetud kaale järgmises korduses. Kuna tunnuste kaalumisel kasutatakse kaalude juhusliku ulatusega muutmist, siis ei ole Constud õpe **determineeritud protsess**, vaid osaliselt juhuslik.

Süsteemis Constud on tunnuste arvutamine (kui juhuvalimit ei kasutata) ja hinnangute arvutamine **determineeritud protsessid**, tehisõpe sisaldab aga juhuslikkust

Näidiste valik. Kasutatakse eelmises etapis saadud parimaid tunnusekaale. Näidiste kaale sobitatakse ükshaaval. Algul on kõigi vaatluste kaal ühtviisi = 1. Seejärel proovitakse parajasti järjekorras olevat vaatlust näidiste hulgast välja jätta ning võrreldakse hinnangu täpsust vaatluse kasutamisel näidisenäidise ja ilma. Jätkatakse selle variandiga, mis annab paremaid tulemusi.

Näidiste kaalumine. Iga vaatluse puhul võrreldakse kolme valikut: suurendada selle vaatluse kaalu näidisenäidisenäidise, vähendada kaalu või jätta see samaks. Täpsemad prognoosid andev variant jääb edasisse kasutusse. Kõigi vaatluste kaalumist korratakse järgmisel kaalumuutuse tasemel. Kaalumuutuste esimene tase on 0,5, iga järgnev tase on eelmisest kaks korda väiksem. Kaalumistasemetete arvu valib kasutaja.

Kui tehisõppe käigus leitakse seni parim tunnuste ja vaatluste kaalude komplekt, siis salvestab Constud need kaalud andmebaasi.

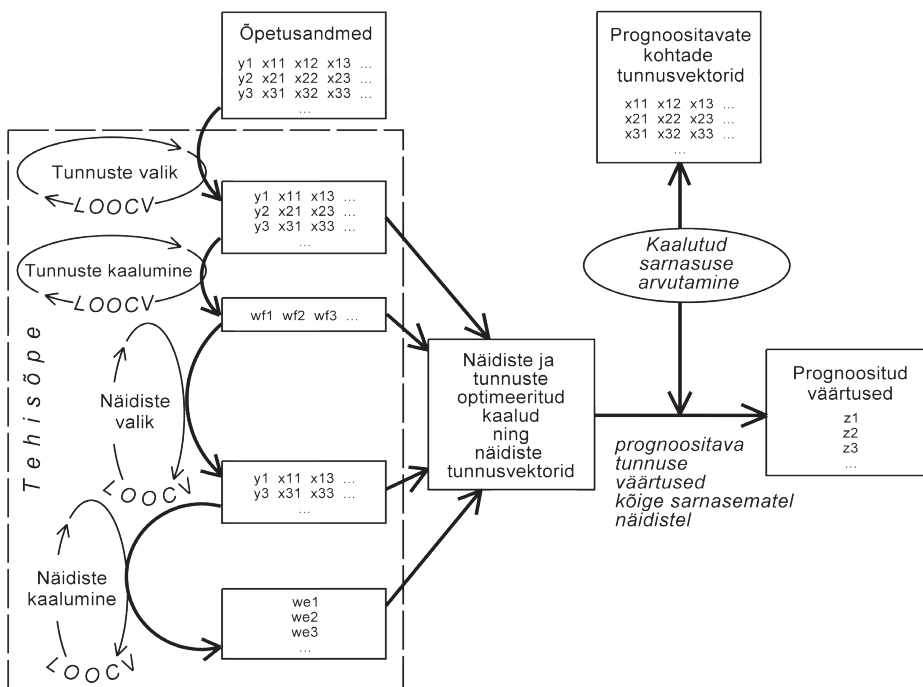
Tunnuste ja näidiste aktuaalsuse muutmine. Aktuaalsusena salvestatakse

õppe käigus järk-järguliselt selgub tunnuste ja vaatluste kasulikkus hinnangute arvutamisel. Kui kasutaja ei ole muud väärtust sisestanud, on aktuaalsuse algväärtus 100. Nende tunnuste ja vaatluste, millest on hinnangute arvutamisel abi, aktuaalsust suurendatakse iga õppevooru järel. Kui tunnust või vaatlust üritati prognoosimisel kasutada, kuid hinnangud on täpsemad ilma selle tunnuse või vaatluseta, siis aktuaalsust vähendatakse. Tunnuste ja näidiste indikaatorväärtus sõltub nende kombinatsioonist. Sama tunnuse kasulikkus teises tunnuste ja vaatluste juhuvalimis võib sootuks erinev olla.

Tunnuste ja vaatluste järgmisesse juhuslikku õppevalimisse kaasamise tõenäosus on aktuaalsusega proportsionaalne. Aktuaalsuse muutmise määra (õppe tundlikkust ehk labiilsust ehk kergeusklikkust) saab seadistada Constud põhiaknas olevatest nihikutest (joonis 4-3). Kui nihik on kõrgemas asendis, siis muutuvad aktuaalsused iga õppevooru järel suuremal määral ja tunnuste aktuaalsused eristuvad kiiremini, kuid juhuslikkuse mõju on suurem.

Aktuaalsused on Constud süsteemi mälu

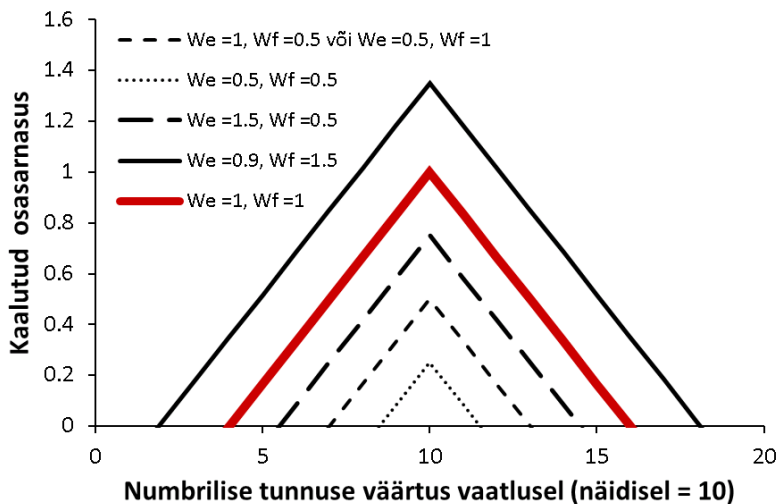
Umbes kümne õppevooru järel normeeritakse tunnuste ja näidiste aktuaalsusi nii, et kasutusel olevate tunnuste ja näidiste keskmine aktuaalsus oleks 100. Tunnuste keskmine kaal hoitakse pidevalt võrdne ühega — kaalude summa võrdub tunnuste arvuga. Näidiste keskmist kaalu ei normeerita.



Joonis 4-1. Tehisõppe ja prognoosi arvutamise tehnoloogiline skeem (Tamm ja

mis on kirjas väljal [FTUNNUSED].[nullkaugus]. Nullkauguse kasutamise eesmärk on vähendada ruumilise autokorrelatsiooni mõju. Ruumilise aspektiga tunnuste puhul on lähestikku tehtud mõõtmistulemused sarnased seetõttu, et enamik nähtusi on mingil määral ruumiliselt pidevad. Selle nähtuse — ruumilise autokorrelatsiooni tõttu ei saa lähestikku tehtud vaatlusi käsitleda sõltumatutena. Nullkaugus tõkestab lähestikuste kohtade vastastikut kasutust prognoositava tunnuse hindamisel. Constud tehisoöpe eelistab hajusamalt paiknevaid ja tänu sellele ruumiliselt esinduslikumaid näidiskohti.

Kõigi õppevalimis olevate vaatluskohtade vahel arvutatakse kauguskorrektuur, mis on nullkauguse ja vaatluskohtade vahemaa jagatis. Kauguskorrektuur lahutatakse osasarnasuste kaalutud keskmisena saadud üldsarnasusest. Kui tulemus on väiksem kui parameeter nullsarnasus, siis omistatakse üldsarnasusele nullsarnasuse väärtus.



Joonis 4-2. Kaalude mõju sarnasusele süsteemis Constud kasutatud sarnasuse arvutamise algoritmi puhul. We — näidise kaal; Wf — tunnuse kaal; jooned kujutavad sarnasust näidisega vaid ühe, horisontaalteljel kujutatud tunnuse järgi. Näidisel on selle tunnuse väärtus = 10 ja selle tunnuse standardhälve = 3.

4.1.2. HINNANGU ARVUTAMINE

Kui sarnasusele tuginevalt järeldatakse vaid ühe näidise järgi, siis on prognoositavaks väärtuseks funktsioontunnuse väärtus sellel ainsal näidisel. Enamasti kasutatakse siiski rohkem kui ühte näidist. Otsustamiseks vajaliku näidiste arvu määrab Constud tarkvaras otsitava sarnasuse hulk. Selle parameetri algväärtuse peab kasutaja sisestama väljale [FTUNNUSED].[sumsimmax]. Kas koos tunnusekaaludega sobitatakse õpetusandmetele ka otsitava sarnasuse hulk, valib kasutaja Constud dialoogiakna tehisoöpe osast (joonis 4-3).

Kõige sarnasemate näidiste leidmine algab sarnasuse tasemelt 1. Kui näidiste

sarnasuste summa sellel sarnasuse tasemel on väiksem kui otsitav sarnasuse hulk, siis alandatakse taset 0,01 võrra. Kõige sarnasemate näidiste otsimine jätkub kuni otsuse langetamiseks otsitav sarnasuse hulk on leitud või kuni näidiseid jätkub. Otsuse langetamisel on näidise mõju proportsionaalne tema sarnasusega. Kui otsitava sarnasuse hulk on suur, siis on hinnangud vähem tundlikud vigade ja juhuslikkuse suhtes õpetusandmetes. Liigsuure otsitava sarnasuse hulga puhul on tulemuseks kõige sagedasema variandi või keskmisele lähedaste väärtuste ülemäära sage pakkumine, mis ka ei pruugi kasutajat rahuldada. Keskvärtus või mood ei anna valdava enamiku prognoositavate vaatluste puhul uut teavet ning nende pakkumiseks pole mõtet kasutada keerukat protseduuri.

4.1.3. TÕESUSKRITEERIUM

Prognoositav väärtus saadakse Constud õpetusandmetes jäta-üks-välja ristkontrolli abil — iga vaatluse puhul arvutatakse prognoositav väärtus kasutades kõiki teisi vaatlusi õpetusvalimis. Kaalude komplektide sobivuse võrdlemisel juhindub Constud tõesuskriteeriumist, mis sõltub prognoositava muutuja tüübist ja mis on täpsustatud peatükkides 4.4...4.8.

4.2. CONSTUD ÕPPE SUVANDID

Kasutaja saab Constud dialoogiakna tehisõppe osast seada järgmisi suvandeid (joonis 4-3).

Ruumiandmed — kui on valitud, siis peavad vaatlusandmete tabelis olema vaatluskohtade koordinaadid ning õppes rakendatakse kauguskorrektuuri.

Otsitava sarnasushulga sobitamine — kui on valitud, siis sobitatakse koos tunnuste kaaludega ka otsitavat sarnasushulka. Kui ei ole valitud, siis kasutatakse väljal [FTUNNUSED].[sumsimmax] olevat sarnasuse summat konstandina.

Standardhälbed tabelist STDEV — kui on valitud, siis otsib Constud arvtunnuste valmisarvutatud standardhälbeid tabelist STDEV. Kui ei ole valitud, siis arvutatakse standardhälbed õpetusvalimist, mis on üldjuhul vaid osa õpetusandmetest ja seetõttu vähem usaldusväärne.

Veateated — kui on valitud, siis saab Constud õppe käigus ilmnenuid tõrgete puhul veateateid saata. Valik mõjutab vaid õppe moodulit, teiste tööde puhul puudub võimalus veateateid alla suruda.

Näidiste kaalumisi — korduste arv näidiste kaalumisel. Võimalikud väärtused on nullist kolmeni. Nulli puhul näidiseid vaid valitakse, aga ei kaaluta. Ühe puhul võivad

kaalud olla 0, 0,5, 1 ja 1,5; kahe puhul: 0; 0,25; 0,5, 0,75; 1; 1,25; 1,5; 1,75; kolme puhul: 0; 0,125; 0,25; 0,375; 0,5; 0,625; 0,75; 0,875; 1; 1,125; 1,25; 1,5; 1,625; 1,75 ja 1,875.

Tunnuste kaalumisi — korduste arv tunnuste kaalumisel. Erinevalt näidiste kaalumisest muudetakse tunnuste kaale juhusliku arvu võrra vahemikus $-0,5...+0,5$.

Lõpeta, kui logiID > — õpe katkestatakse, kui **logitabelisse** salvestatud kirje ID number on suurem kui valitud väärtus. Tänu sellele **suvandile** saab planeerida ja ellu viia **tehisõppe** katseid, mille tulemused on tänu samale õppevoorude arvule omavahel võrreldavad.

Nullsarnasus — murdarvuline sarnasuse tase, millest väiksemad sarnasuse väärtused tasandatakse nullsarnasuse tasemele. Üldjuhul ei ole tarvidust selle suvandi väärtust nullist erinevaks seada. Väärtuste ≥ 1 korral lakkab sarnasuste eristamine, negatiivsed väärtused võimaldavad hinnangutesse kaasata näidiseid, mis erinevad arvtunnuste poolest enam kui üldjuhul kaasatavad väärtused.

Sarnasuse ulatus SD — arvtunnuse väärtuste erinevus standardhälvetes, millest suurema ja võrdse erinevuse korral loetakse **osasarnasus** selle tunnuse osas nulliks. Sarnasuse ulatus mõjutab sarnasuse ulatuse laiust arvtunnuse teljel (joonis 4-4).

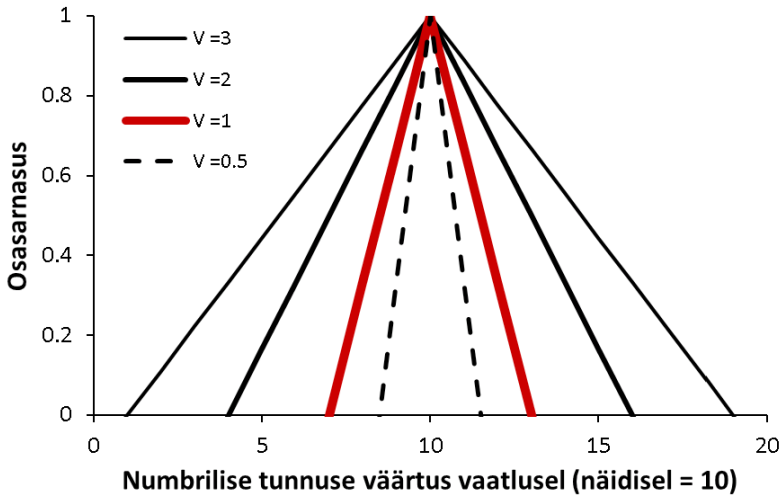
Vaatluste suhtes kergeusklik/umbusklik — nihiku asend mõjutab **vaatluse** aktuaalsuste muutmise määra, mis omakorda määrab vaatluse õppevalimisse sattumise tõenäosuse.

Tunnuste suhtes kergeusklik/umbusklik — nihiku asend mõjutab tunnuse aktuaalsuste muutmise määra, mis omakorda määrab tunnuse õppevalimisse sattumise tõenäosuse.

Aktuaalsused on Constud tehisõppesüsteemi mälu ning mõõdavad tunnuste indikaatorväärtust. Aktuaalsemad vaatlused ja tunnused on eelnenud õppe käigus enamasti prognoosi arvutamisel abiks olnud, väheaktuaalsed tunnused ja vaatlused on hinnangute arvutamisel enamasti kasutatud olnud. Aktuaalsused määravad tunnuste ja vaatluste õppevalimisse sattumise tõenäosust.

Indeksite arvutus	Tehisõpe	Teadmiste kontroll	Hinnangute arvutus
<input checked="" type="checkbox"/> Ruumiandmed	Vaatluste kaalumisi	1	Vaatluste suhtes
<input checked="" type="checkbox"/> Otsitava sarnasushulga sobitamine	Tunnuste kaalumisi	20	kergeusklik
<input type="checkbox"/> Standardhälbed tabelist STDEV	Lõpeta, kui logi ID >	1000	umbusklik
<input type="button" value="Arvuta standardhälbed"/>	Nullsarnasus	0	Tunnuste suhtes
	Sarnasuse ulatus [SD]	2	kergeusklik
<input checked="" type="checkbox"/> Veateated			umbusklik

Joonis 4-3. Constud tehisõppe seadistamise aken.



Joonis 4-4. Sarnasuse ulatuse mõju sarnasuse väärtusele.

Constud õpe ei jõua samade kaalude ja sama hinnangutäpsuse juurde ka siis, kui õpet korrata täpselt samade andmete ja **suvanditega**, sest tunnuste parimaid kaale otsitakse neid juhuslikul määral nihutades. Juhulik nihe on iga kord erineva suurusega.

Constud õpet saab katkestada:

- 1) klikkides x-märki Constud protsessi edenemise akna parempoolses ülanurgas,
- 2) Windows tegumihaldurist (Task Manager nupuga *End process* või nupuga *End task*),
- 3) lülitades välja kõik read väljal *[FTUNNUSED].[arvutada]* (õpe lõpeb poolelioleva õppevooru lõpetamise järel).

Õppe senine seis säilib Constud andmebaasis, kuna tulemused salvestatakse iga õppevooru järel. Õppe taaskäivitamisel jätkub see poolelijäänud seisust.

Constud õpet võib katkestada, õppe tulemused salvestuvad iga õppevooru järel

4.3. FUNKTSIOONTUNNUSE EELKLASSIFIKAATOR

Funktsioontunnuse eelklassifikaator on üks argumenttunnuste tabelis määratletud nominaalsetest tunnustest, mille AID number on väljal *[FTUNNUSED].[eelklass]* ja mille number on suurem kui null. Eelklassifikaatori klassid piiravad näidiste

kasutamist vaid selle ühe klassi piires. Funktsioontunnuse eelklassifikaator mõjutab nii õpet kui ka hinnangute arvutamist. Kui eelklassifikaatori kasutamisel valib **Constud tehisõpe** iga eelklassi jaoks eraldi näidised, siis sarnasuse arvutamisel kasutatavad tunnused ja nende kaalud on kõigis eelklassides samad.

Igas eelklassis on eraldi näidised, tunnuste kaalud on eelklassides samad

4.4. KAHEVÄÄRTUSELINE MUUTUJA

Kaheväärtuselisel ehk binaarsel ehk Boole'i muutujal on vaid kaks võimalikku väärtust, näiteks jah või ei, null või üks, pluss või miinus, õige või vale.

Binaarse muutuja hinnangute **vastavust** tõeseks loetavatele väärtustele mõõdetakse **Constud** programmis Hanssen-Kuiper skoori järgi (*true skill statistic — TSS*), milles loendite asemel kasutatakse sarnasuste summasid. TSS eeliseks on nii tõese positiivse kui ka tõese negatiivse hinnangu osakaalu võrdne arvestamine sõltumata positiivsete ja negatiivsete juhtude sagedusest **õpetusandmetes**. Muutumisvahemiku poolest on TSS nagu **Pearsoni korrelatsioonikordaja**, omades võimalikke väärtusi miinus ühest pluss üheni ning null märkimas juhuslikkuse korral oodatavat väärtust. Juhuslikult kahe võimaluse vahel valides tabame õige variandi pooltel positiivsetel ja pooltel negatiivsetel juhtudel. Kui kahest variandist on teadaolevalt üks palju sagedasem kui teine, siis võib õigete hinnangute kõrge osakaalu saavutada pakkudes alati seda teadolevalt sagedasemat varianti. Näiteks, kui positiivseid juhtumeid on vaid 1% vaatlustest, siis hinnates kõik **vaatlused** negatiivseteks, saame väita, et hinnangute kokkulangevus tegelikkusega on 99%. TSS statistiku järgi vastab hinnangute täpsus sellisel juhul siiski vaid juhuslikkuse korral oodatavale väärtusele null (*Allouche et al., 2006*).

Modifitseeritud TSS arvutamisel **Constud** tarkvaras liidetakse õigestihinnatud positiivsete juhtude osakaalu ja õigestihinnatud negatiivsete juhtude osakaalu asemel õigesti hinnatud sarnasuse osakaalud, muutumisvahemiku tavapäraseks nihutamiseks lahutatakse üks.

$$MTSS = TPS + TNS - 1$$

MTSS — tõepäraskoor sarnasuste järgi,

TPS — õigesti hinnatud sarnasuse osakaal positiivsete juhtudega,

TNS — õigesti hinnatud sarnasuse osakaal negatiivsete juhtudega.

Constud andmebaasi väljad, kus hoitakse binaarse muutuja vaadeldud ja prognoositud väärtusi, võivad olla kas binaarses (*Ei/Jah*) vormingus või ühebaidise arvuna. Baidivormingu puhul tuleks kasutada arvkoode null ja üks.

Enne õppe käivitamist tuleb Constud andmebaas vajalikul kujul ette valmistada.

42. Lülita välja eelmiste muutujate arvutus.

Uuritavate muutujate valikut saab seada väljalt *[FTUNNUSED].[arvutada]* kõigi tööde puhul programmiga Constud.

43. Lisa funktsioontunnuste tabelisse uus prognoositav muutuja soo-neiuvaiba (*Epipactis palustris*) esinemise või puudumise käsitlemiseks.

Võid kopeerida tabelis *FTUNNUSED* oleva rea ja muuta õppevalimi mahuks 100, kontrollivalimi mahuks 500, seletavate tunnuste valimi mahuks 6, null-kauguseks 100 ja muutuja tüübi koodiks *ftyyp* = 3 (vt ka peatükk 2.2.1 ja lisa 2.1). Uue muutuja atribuudid on kirjas ka alltoodud päringus.

```
INSERT INTO [FTUNNUSED]
([arvutada], [FID], [antudkood], [logiID], [eelklass], [nimi], [ftyyp],
[osatunnuseid], [sumsimmax], [andmepäring],
[vmah], [kvmah], [tmah], [nullkaugus])
VALUES (True, 2, 0, 0, 0, 'E. palustris on/pole', 3, 0, 3,
'VAATLUSKOHAD1', 100, 500, 6, 100);
```

44. Lisa tabelisse *ATUNNUSED* väljad *F2* ja *AF2*, mis vastavad prognoositavale muutujale *FID* numbriga 2.

Saad kasutada nuppe *Väljad E... ja AF...* andmetabelite loomise aknas või kopeerida ja ümber nimetada olemasolevad väljad *F1* ja *AF1* või kasutada allolevat päringut.

```
ALTER TABLE [ATUNNUSED] ADD [F2] bit NULL, [AF2] byte NULL;
```

45. Lisa tabel *LOG_F2*, mis vastab uuritavale muutujale *FID* = 2.

Constud tabelite loomise akna nupp *Logitabelid* loob logitabelid kõigi muutujate jaoks, mis on kirjas tabelis *FTUNNUSED*.

46. Lülita sisse kõigi seletavate tunnuste kasutus väljal *[ATUNNUSED].[F2]*, välja arvatud tunnus 10.

Välja *[ATUNNUSED].[F1]* ei ole tarvis muuta. Tunnus 10 jääb välja, sest seda ei ole õpetusandmete tabelis. Tunnuse 10 arvutasime vaid rasterfaili. Soo-neiuvaiba vaatluskohad on juba tabelis *VAATLUSKOHAD1*.

```
UPDATE [ATUNNUSED] SET [F2] = True WHERE AID<>10;
```

Väljal F... võivad sisse lülitatud olla vaid need tunnused,
mis on õpetusandmetes tõesti olemas

47. Nüüd ja alati enne Constud õppe algust — kontrolli andmebaasis järgmist.



1. Kas väljal [FTUNNUSED].[arvutada] on valitud vajalikud read?
2. Kas valitud funktsioontunnusele vastavad väljad [F...] ja [AF...] tabelis [ATUNNUSED] on olemas?
3. Kas väljal [ATUNNUSED].[F...] on valitud vajalikud tunnused?
4. Kas õpetusandmete tabelis või päringus (praegusel juhul VAATLUSKOHAD1) on olemas väljad VID, F, progn, prob, w, alates, kuni ning seletavate tunnuste väljad, mis on väljal [ATUNNUSED].[F...] sisse lülitatud?
5. Kas on olemas uuritavate muutujate logitabelid ja nendes vähemalt üks olemasolev rida (joonis 4-5)?


FTUNNUSED		
FID	antudkood	
LOGI_F1	1	
LOGI_F2	2	
LOG_F3	3	
LOGI_F4	4	
LOGI_F5	5	2
	5	46
	5	68
	5	96
	5	103
LOGI_F6	6	

Joonis 4-5. Seos uuritavate muutujate ja logitabelite vahel.

Access andmebaasi kasutamisel veendu ka, ega objekt ei ole avatud disainvaates, sest disainvaates avatud **andmebaasiobjekt** on muu kasutuse jaoks lukustatud.

Constud andmebaasi struktuuri saab kontrollida nupuga *Tabelite kontroll* (joonis 2-2). Andmebaasstruktuuri formaalne kontroll ei taga kirjete sisulist õigsust.

- Enne Constud õpet kontrolli:
- 1) [FTUNNUSED].[arvutada],
 - 2) F... ja AF... tabelis ATUNNUSED,
 - 3) väljad VID, F, progn, prob, w, alates, kuni ning seletavate tunnuste väljad,
 - 4) logitabel,
 - 5) kujundusvaated kinni

 **48. Käivita Constud ja vali tehisõppe osa. Säilita vaikumisi seaded (joonis 4-3), välja arvatud õppevoorude arv (Lõpeta, kui logi ID >), kus õppeharjutuse puhul peaks kahesajast piisama. Käivita tehisõppe vajutades nupule Jätka.**

200 õppevooru valimitega, kus on ligikaudu sada vaatlust ja kuus tunnust, võtab heas arvutis keskmiselt 10 minutit aega. Õppe edenemist saab õppe ajal jälgida logitabelist. Samuti on õppe ajal võimalik muuta neid õppe parameetreid, mis on andmebaasis. Enamiku uurimisülesannete jaoks piisab tuhandest kuni paarist tuhandest voorust. Parima prognoosiva komplekti otsimisel võiks suurema arvu õppekorduste asemel eelistada paralleelseid õppeid.

 **49. Ava logitabel LOG_F2 ja leia õppevoor, mille kontrollvastavus on suurim.**

Pane tähele, et sama ID number on kirjas väljal [FTUNNUSED].[logiid]. Peale selle on õppe tulemusena salvestunud: vastavus õppevalimis — [LOG_F2].[tvastavus], vastavus kontrollvalimis — [LOG_F2].[kvastavus], näidiste arv — [LOG_F2].[etn], õppevalimi suurus — [LOG_F2].[valim], seletavate tunnuste kaalud — [LOG_F2].[tunnusekaalud], seletavate tunnuste aktuaalsused — [ATUNNUSED].[AF2], vaatluskohtade aktuaalsused — [VAATLUSKOHAD1].[aktuaalsus], näidiste kaalud parimas komplektis — [VAATLUSKOHAD1].[w], parima tulemusega õppevooru kontrollvalimis olevatele vaatlustele prognoositud väärtused — [VAATLUSKOHAD1].[progn], prognoositud koha sarnasus prognoosimisel kasutatud näidistega — [VAATLUSKOHAD1].[prob].

Vaatluste kaalud ja aktuaalsused arvutatakse
õppevalimis, prognoositud väärtused ja sarnasus aga
kontrollvalimis

Pane tähele järgmist.

1. Tunnuste kaalud on logitabelis kujul: <tunnuse AID number> <selle tunnuse kaal> <tunnuse AID number> <selle tunnuse kaal> jne.

2. Aktuaalsuste keskmine on saja lähedal, sest **Constud** aeg-ajalt normeerib aktuaalsuskoodi.
3. Õppes moodustatud valimite suurus ei ole täpselt see, mis väljal *[FTUNNUSED].[valim]* ette antud, sest **Constud** kasutab iga **vaatluse** valimisse kaasamise või valimist välja jätmise otsustamisel kaasamistöenäosust ja ei kohanda **valimi mahtu** täpselt etteantud suuruseni.
4. Vaatluste kaalud ja aktuaalsused saadakse õppevalimis, prognoositud väärtused ja sarnasus **näidistega** salvestatakse aga kontrollvalimis olevatele vaatlustele.
5. Suurimate aktuaalsuskoodidega vaatluskohad ei ole alati suurima kaaluga — aktuaalsused näitavad indikaatorväärtust kumulatiivselt; kaalud on vaid ühe õppevooru tulemus — selle vooru, mille **kontrollvastavus** on seni parim.

Aktuaalsused näitavad indikaatorväärtust, kaalud seni parimat komplekti

4.5. MULTINOMIAALNE MUUTUJA

Multinomiaalne muutuja on nominaalne ehk kategooriline muutuja, millel on rohkem kui kaks võimalikku väärtust. Multinomiaalse muutuja hinnangute ja tõeseks loetavate tulemustevahelise vastavuse määra mõõdetakse klassifikatsioonivastavuse kordajaga ehk kapa kordajaga (*Congalton ja Green, 1999*). Kapa kordaja teoreetiline muutumisvahemik on miinus ühest pluss üheni, null vastab juhuslikkuse korral oodatavale.


Nädisandmete andmebaasi tabelis *Taimkate* on taimkatte detailse kaardistamise tulemused Eesti põhikaardi lehelt 54441. Kaardistusel üritati eristada 111 taimkatteüksust. Tabel on ette valmistatud vastavalt **Constud** vajadustele. Vaatluskohtade tunnuseid ei ole seekord tarvis arvutada, sest harjutamiseks vajalik tunnuste hulk on tabelis olemas.

Enne **Constud** õpet tuleks:

- 1) lisada taimkatteüksuste õppe rida **funktsioontunnuste** tabelisse,

- 2) lisada funktsioontunnuse FID numbrile vastavad väljad *F...* ja *AF...*,
- 3) lisada tunnuste määratlused tabelisse *ATUNNUSED*,
- 4) lisada **logitabel**,
- 5) valida funktsioontunnuse ja kasutatavate **argumenttunnuste** kirjed.


 **50. Impordi taimkatteandmete tabel Taimkate andmebaasist Constud_6ppe_andmed.mdb oma Constud andmebaasi.**

 **51. Lisa tunnuste 13...50 määratlused näidisandmete andmebaasi tabelist Lisatunnused Constud andmebaasi tabelisse ATUNNUSED.**

Tegelikult võiks valmis arvutatud tunnused siduda ükskõik millise andmekihi nominaalse või numbrilise indeksiga, sest õppes kasutatakse seletavate tunnuste **atribuutidest** vaid indeksi tüüpi, mis määrab tunnuse tüübi — nominaalne või numbriline (vt ka peatükk 2.1).

 **52. Kopeeri täiendavate andmekihtide 12...33 kirjeldused näidisandmete andmebaasi tabelist Lisakihid tabelisse AKIHID Constud andmebaasis.**

Tegelikult võiks valmis arvutatud tunnused siduda ükskõik millise andmekihiga, mille kasutusaeg ei ole piiratud, sest andmekihti vajatakse Constud õppes vaid tunnuste kasutuspiirangute määramiseks. Universaalse nominaalse ja universaalse numbrilise andmekihi kasutamist mitteruumilise muutuja puhul näidatakse peatükis 4.8. Andmekihtide ja tunnuste seostest on juttu peatükis 2.1.

 **53. Moodusta andmebaasivaade, mis valiks tabeli Taimkate kõigi väljade kirjed, kus VID < 301.**

Anna vaatele nimi, näiteks *Taimkatteõpe*. Järgmises peatükis üritame tuvastada sellest õpetusandmete valimist välja jäänud vaatluste taimkatteüksust.

```
SELECT * FROM [Taimkate] WHERE VID<301;
```

 **54. Lisa funktsioontunnuste tabelisse rida, mis on kirjeldatud järgmises päringus.**

Veendu, et väljale *andmepäring* kantaks ülesandes 53 moodustatud päringu nimi. Pane tähele, et *[FTUNNUSED].[ftyyp]* = 0 tähistab **multinomiaalset muutujat** (lisa 2.1 ja peatükk 2.2.1), tunnus 13 on seekord **funktsioontunnuse eelklassifikaatoriks**, valimi suurus on nii õppes kui ka kontrollis 150; ja lõpuks — **õpetusandmeid** ei loeta otse andmetabelist, vaid andmebaasivaate kaudu.

```
INSERT INTO [FTUNNUSED] ([arvutada], [FID], [antudkood], [logiID],
[eelklass], [nimi], [ftyyp], [osatunnuseid], [sumsimmax], [andmepäring],
[vmaht], [kvmaht], [tmaht], [nullkaugus])
VALUES (True, 3, 0, 0, 13, '111 taimkatteüksust', 0, 0, 2,
'Taimkatteõpe', 150, 150, 10, 100);
```

55. Lülita välja varasemate funktsioontunnuste arvutus väljal [FTUNNUSED]. [arvutada]. 

56. Lisa tabelisse ATUNNUSED väljad [F3] and [AF3]. 

Ülesannet 44 võib eeskujuks võtta. Kontrolli, et lisatavate tunnuste numbriline osa vastaks funktsioontunnuse FID numbrile, seda juhendit järgides on see number 3.

57. Uuenda [ATUNNUSED].[F3] väärtuseks True kirjetel, mille AID > 12 ja False kirjetel, mille AID < 13. 




SQL päringutena:

```
UPDATE [ATUNNUSED] SET [F3] = True WHERE [AID]>12;
```

```
UPDATE [ATUNNUSED] SET [F3] = False WHERE [AID]<13;
```

58. Lisa funktsioontunnuse logitabel LOGI_F3. 

Soovitame kasutada Constud tabelite loomise funktsiooni. See teeb puuduvad logitabelid kõigile tabelis FTUNNUSED loetletud muutujatele.


59. Kontrolli ülesandes 47 loetletud andmestruktuure ja seejärel käivita Constud õpe. Säilita vaikumisi suvandid (joonis 4-3) välja arvatud õppekorduste arv, mille puhul esmatutvusel piisab kahesajast.   

Õppe tulemuse salvestamiskohad on kirjas ülesande 49 juures. Multinomiaalse muutuja ainus erinevus kaheväärtuselisest on muutuja ja selle hinnangute ühebidine vorming.

4.6. NUMBRILINE MUUTUJA

Numbrilist muutujat käsitleb Constud pideva murdarvulise muutujana. Samad arvutusreeglid on rakendatavad ka diskreetse järjestatava muutuja korral, kui selle väärtused on esitatud neljabaidiste murdarvudena. Numbrilise muutuja hinnangutäpsust määrab Constud ruutkeskmise vea järgi, mille arvutamisel on vaatluste arvust lahutatud kasutatud tunnuste arv, hindamaks vabadusastmete arvu.

Näidisandmete andmebaasi on valmis seatud tabel *Baltisademed*, kus on Baltimaade vaatlusjaamades pikaajaliselt määratud aasta keskmine sademete hulk millimeetrites ning mõned jaama asukohta ja selle ümbrust kirjeldavad tunnused. Väljad *VID, x, y, alates, kuni, F* (kolmekümne aasta keskmine sademete hulk), *progn, prob, aktuaalsus, w* ning numbrilise nimega väljad seletavate tunnuste jaoks vastavad Constud andmestruktuurile.

→  **60. Impordi või kopeeri tabel *Baltisademed* andmebaasist *Constud_6ppe_andmed.mdb* Constud andmebaasi.**

→  **61. Kopeeri Constud andmebaasi tabelisse *ATUNNUSED* tunnuste *51...57* määratlused näidisandmete andmebaasi tabelist *Lisatunnused*.**

Numbritega 51...57 tähistatud tunnused on järgmised.

51 — Lambert-Est süsteemi lääne-ida suuna ristkoordinaat jagatud 5000 (ühebaidisesse vormingusse teisendamiseks).

52 — Lambert -Est süsteemi lõuna-põhja suuna ristkoordinaat jagatud 10 000 (ühebaidisesse vormingusse teisendamiseks).

53 — 75 km raadiuses olevates jaamades mõõdetud pikaajaline aasta sademete hulk sentimeetrites, sest seletavad tunnused peavad olema ühebaidises vormingus.


54 — veekogude osa maakattes 10 km ümbruses.

55 — metsa osa maakattes 10 km ümbruses.

56 — kaugus mereni kilomeetrites.

57 — maapinna kõrgus meetrites.

Nende kirjeldavate tunnuste ainsad õppes vajalikud atribuudid on andmekihi number *KID* ja indeksi number *indeksID*, kuna tunnuste väärtused on juba valmis arvutatud (peatükk 2.1).


 **62. Lisa tabelisse *AKIHID* universaalse nominaalse (*KID = 10*) ja universaalse numbrilise andmekihi (*KID = 11*) atribuudid.**

Eelmises ülesandes lisatud tunnused on juba seotud universaalse numbrilise kihiga. Universaalset nominaalset kihti läheb tarvis edaspidi. Need kihid sobivad ühiseks kasutuseks varem valmis arvutatud tunnuste ja mitteruumiliste muutujate puhul. Constud vajab, et iga tunnus oleks mingi andmekihiga seotud (peatükk 1.2).

Ülesandes mainitud kahe kihi määratlused saab kopeerida näidisandmete andmebaasi tabelist *Lisakihid* või siis lisada alltoodud päringute abil.

```
INSERT INTO [AKIHID] ([KID],[nominal],[kataloog])
VALUES (10,True,'universal nominal');
```


```
INSERT INTO [AKIHID] ([KID],[nominal],[kataloog])
VALUES (11,False,'universal numerical');
```


63. Lülita väljalt [FTUNNUSED].[arvutada] välja varasemate funktsioontunnuste arvutus. 

64. Lisa järgmise päringuga määratletud rida tabelisse FTUNNUSED. 

Kontrolli, kas väljale *andmepäring* kantakse eelnevalt imporditud tabeli nimi. Pane tähele, et [FTUNNUSED].[ftyyp] = 2 tähistab numbrilist muutujat (lisa 2.1 ja peatükk 2.2.1) ja et nii õppe- kui ka kontrollvalim mahuga 123 vaatluskohta hõlmab kogu õpetusandmestikku.



```
INSERT INTO [FTUNNUSED]
([arvutada],[FID],[nimi],[ftyyp],[sumsimmax],[andmepäring],
[vmaht],[kvmaht],[tmaht],[nullkaugus]) VALUES
(True,4,'Baltic precipitation',2,3,'Baltisademed',123, 123,6,100);
```

65. Lisa tabelisse ATUNNUSED väljad [F4] ja [AF4] ja uuenda [ATUNNUSED].[F4] = True seal, kus AID > 50. 

Teised selle välja kirjed on vaikimisi = False. Juhised on ülesande 44 juures. Kontrolli, et väljade F... ja AF... numbriline osa vastaks funktsioontunnuse FID numbrile.

66. Lisa selle funktsioontunnuse logitabel. 

Constud tabelite loomise akna nupp *Logitabelid* loob logitabelid kõigi muutujate jaoks, mis on kirjas tabelis FTUNNUSED.

67. Käivita Constud õpe. Säilita vaikimisi suvandid (joonis 4-3) välja arvatud korduste arv, mille puhul piisab esialgu kahesajast.  

Õppe tulemused salvestuvad andmebaasi väljadele: AF, w, aktuaalsus, progn, prob, tvastavus, kvastavus, tunnusekaalud jne, nagu eelmiste muutujate korralgi. Pane tähele, et hinnangu vastavust vaatlustele mõõdetakse numbrilise muutuja puhul ruutkeskmise hälbega — mida väiksem hälve, seda parem tulemus.

4.7. ÜKS MULTINOMIAALSE MUUTUJA ÜKSUS

Constud võimaldab õppida **nominaalse muutuja** üksiku kategooria ehk üksuse tuvastamist sama muutuja kõigi ülejäänud klasside ühendhulgast. Seega saab parimad kaalud salvestada iga klassi jaoks eraldi. Arvutuslikult käsitletakse üksiku klassi eristamise ülesannet nagu binaarse muutuja õpet. Ka **tõesuskriteerium** on sama — modifitseeritud TSS (peatükk 4.4). Järgnevalt kasutame küll sama taimkatte andmestikku, mida peatükis 4.5, kuid mõned andmete ettevalmistamise tööd tuleb siiski enne õpet ära teha.



68. Loo tabeli kujundusvaates koopia õpetusandmete tabeli Taimkate väljadest: *progn*, *prob*, *aktuaalsus*, *w* ja anna uutele väljadele nimed: *progn1*, *prob1*, *aktuaalsus1*, *w1*.

Neid välju läheb tarvis järgmises päringus. Eelmistel väljadel olevad väärtused ei peaks kopeeritud väljadele kaasa tulema. Väli *w1* peaks olema lühikese murdarvu (*single* või *real*, sõltuvalt andmebaasitarkvarast) vormis, ülejäänud kolm välja ühebaitsised täisarvud (*byte*).



69. Loo andmebaasivaade Taimkatteõpe1, mis loeb 300 esimese vaatluskoha andmed tabelist Taimkate ja nimetab ümber väljad: *progn1*, *prob1*, *aktuaalsus1*, *w1* andes neile Constud eestikeelse andmebaasi struktuurile vastavad nimed: *progn*, *prob*, *aktuaalsus*, *w*.

See on näidis, kuidas saab sama tabeli alusel luua mitu õpetusandmete vaadet. Samade nimedega väljad: *progn*, *prob*, *aktuaalsus*, *w* on vajalikud mõlemale õpitavale muutujale. Kui mõlema muutuja õppe tulemused oleksid suunatud füüsiliselt samadele väljadele, kirjutaks üks õpe teise õppe tulemused üle. Samast tabelist andmeid kasutavate õpete tulemuste lahus hoidmiseks lisasime tabelisse kordusveerud ja muutsime nende nimesid vaates, mille nimi on kantud funktsioontunnuste tabelisse. Seega Constud tarkvara jaoks vajalike rangelt ette antud nimedega andmebaasiobjekte saab luua, muutes algseid nimesid andmebaasivaates.

Väljade nimesid muutev SQL käsk on järgmine.

```
SELECT [VID], [F],[X],[Y],[progn1] AS progn, [prob1] AS prob, [w1]
AS w, [aktuaalsus1] AS aktuaalsus,[alates], [kuni],[13],[14],[15],
[16],[17],[18],[19],[20],[21],[22],[23],[24],[25],[26],[27],[28],
[29],[30],[31],[32],[33],[34],[35],[36],[37],[38],[39],[40],
[41],[42],[43],[44],[45],[46],[47],[48],[49],[50]
FROM [Taimkate] WHERE [VID]<301;
```

Salvesta vaade nimega *Taimkatteõpe1*.

Andmebaasiobjektide nimesid saab vaates muuta

- 70. Lülita välja kõigi olemasolevate funktsioontunnuste õpe väljalt [FTUNNUSED]. [arvutada].
- 71. Lisa funktsioontunnuste tabelisse rida iga eraldi õpitava üksuse jaoks.

Võtke esialgu õppida vaid viis suhteliselt sagedast taimkatteüksust: palumännik (kood 2), madalsookaasik (46), aruniit (68), sööt (96) ja künnipõld (103). Palumänniku

lisamise SQL keele käsk on järgnev. Pane tähele, et *ftyyp* = 4 tähistab **nominaalse muutuja** üksiku üksuse õpet ja et väljal *andmepäring* on ülesandes 69 loodud andmebaasivaate nimi.

```
INSERT INTO [FTUNNUSED]
([arvutada],[FID],[antudkood],[logiID],[eelklass],[nimi],[ftyyp],
[sumsimmax],[andmepäring],[vmaht],[kvmaht],[tmaht],[nullkaugus])
VALUES (True,5,2,0,13,'palumännik',4,2,
'Taimkatteõpe 1',150,150,10,100);
```

Enne, kui kasutad seda SQL käsku kirjete lisamiseks teiste taimkatteüksuste jaoks, muuda päringus antud klassi kood ja nimi. Selle ülesande lõpuks peaks **funktsioontunnuste** tabelis olema viis rida, millel FID = 5 ja mis erinevad vaid antud koodi ja nime poolest. Antud koodid esindavad taimkatteüksusi, mille eristamist kaugseire andmete järgi kavatakse õppida.

72. Lisa tabelisse ATUNNUSED väljad [F5_...] ja [AF5_...].



Need väljad on vajalikud iga eraldi õpitava üksuse jaoks, sest tunnused ja tunnuste aktuaalsused on iga õpitava üksuse puhul sõltumatud teiste õpitavate üksuste eristamise tunnustest. Väljade *F...* ja *AF...* lisamiseks saad kasutada **Constud** tabelite loomise vormi, kopeerida ja kohandada olemasolevaid *F...* ja *AF...* välju lisades nende numbrilisele osale alumise kriipsu '_' ja eristatava üksuse koodi või siis käivitada allpool toodud SQL käsud.

```
ALTER TABLE [ATUNNUSED] ADD [F5_2] bit NULL, [AF5_2] byte NULL;
ALTER TABLE [ATUNNUSED] ADD [F5_46] bit NULL, [AF5_46] byte NULL;
ALTER TABLE [ATUNNUSED] ADD [F5_68] bit NULL, [AF5_68] byte NULL;
ALTER TABLE [ATUNNUSED] ADD [F5_96] bit NULL, [AF5_96] byte NULL;
ALTER TABLE [ATUNNUSED] ADD [F5_103] bit NULL, [AF5_103] byte NULL;
```

73. Lisa funktsioontunnusele FID = 5 vastav tabel LOGI_F5.

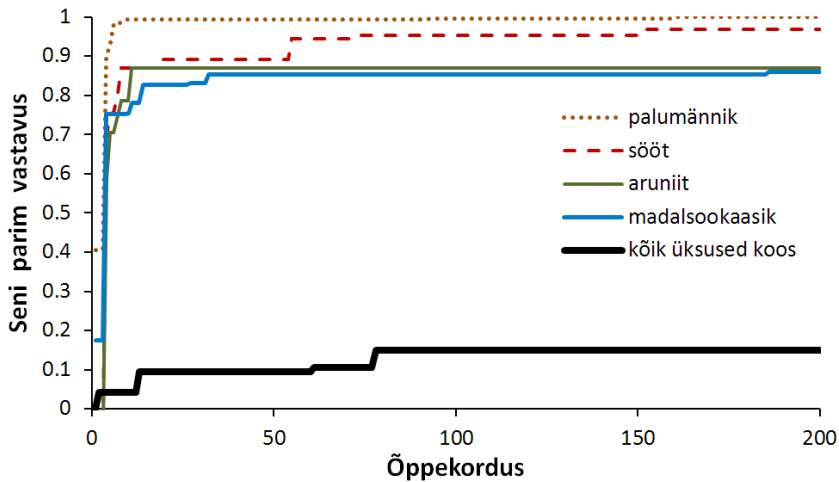


Soovitame kasutada **logitabeli** loomise nuppu **Constud** tabelite loomise vormil, mis loob logitabelid kõigi muutujate jaoks, mis on kirjas tabelis *FTUNNUSED*. Ühise FID numbriga **nominaalse muutuja** kõigi eraldi õpitavate üksuste õppe tulemused salvestatakse samasse logitabelisse. Eri üksuste õppe tulemusi võimaldab eristada logitabeli väli *antudkood*.

74. Lülita sisse seletavate tunnuste AID > 13 ja AID < 51 kasutus veergudes, mis on seotud funktsioontunnusega 5.



```
UPDATE [ATUNNUSED] SET [F5_2]=True,[F5_46] =True,
[F5_68]=True,[F5_96]=True, [F5_103]=True
WHERE [AID]>12 and [AID]<51;
```

Joonis 4-6. Taimkatteüksuste õppekõverate näidis.

4.8. MITTE-RUUMILINE MUUTUJA

Ruumilise muutuja vaatluste kirjeldused sisaldavad asukoha koordinaate, mitte-ruumilise muutuja vaatlused ei ole asukohakoordinaatidega seotud. Constud süsteemis töötlemiseks sobivad mitte-ruumilised andmed on näiteks küsitluse tulemused, kus vastused ja küsitletud isikutega seotud tunnused on kodeeritud täisarvudeks vahemikus 0...254 (255 tähistab andmete puudumist). Mitte-ruumilisi vaatlusi kirjeldavaid tunnuseid ei saa arvutada ruumilistest andmekihtidest, tunnuste väärtused peavad tulema mujalt.

Mitte-ruumilisi tunnuseid ei saa arvutada ruumilistest
andmekihtidest

Mitte-ruumilise muutuja õppeks ja prognoosimiseks ette valmistatud Constud andmebaas erineb ühe olulise pisiasja poolest — **seletava tunnuste** tüüp (numbriline või nominaalne) tuletatakse andmekihi tüübist, mitte indeksi tüübist, nagu oli ruumiliste tunnuste puhul (peatükk 2.1). Vähem oluline erinevus on, et mitte-ruumilise tunnuse uurimisel ei ole tabelid *INDEKSID* ja *PK20t* vajalikud.

Õpetusandmete andmebaasis on tabel *Enterobiasis*, mis sisaldab väljavõtet Kagu-Eesti lasteaedades läbi viidud enterobiaasi esinemise uuringu tulemustest (*Remm, 2005*). Enterobiaas on nakatus pisikese soolenugilise — naaskelsabaga (*Enterobius vermicularis*). Tabeli iga rida esindab ühte last. Väljal *F* on nakatus, igat last kirjeldas algses uuringus palju tunnuseid. Näidisandmetesse on võetud järgmised


tunnused: [60] — vanus, [61] — sugu, [62] — laste arv peres, [63] — ühevanuseline rühm või segavanuseline rühm, [64] — laste vanuse haare lasteaiarühmas, [65] — geograafiline piirkond (maakond või suurem asula), [66] — asula tüüp. Tunnused 61, 65 ja 66 on nominaalsed, ülejäänud nimetatutest numbrilised. Järgnevates ülesannetes püüame nende seitsme tunnuse järgi ära arvata, kas üks või teine laps on nakatatud või mitte ning võrdleme tunnuste indikaatorväärtust. Üksikisiku tasemel hinnangutest on võimalik tuletada nakkuse sagedust piirkonniti ja elanikkonna muudes alajaotustes.

Pane tähele, et asukohakoordinaatideta vaatluste kasutamisel tuleb eemaldada linnuke Constud põhiakna õppepaneeli lahtrist Ruumiandmed.

Enne Constud käivitamist tuleb tegeleda andmete ettevalmistamisega.


 **76. Impordi või kopeeri tabel Enterobiasis näidisandmete andmebaasist Constud andmebaasi.**

 **77. Lülita väljalt [FTUNNUSED].[arvutada] välja kõigi olemasolevate funktsioontunnuste arvutus.**

 **78. Lisa uus funktsioontunnus enterobiaasi nakatunute ja tervete laste eristamise õppeks ja prognoosimiseks.**

See uus muutuja on kaheväärtuseline ja tuleks väljal [FTUNNUSED].[ftyyp] kodeerida arvuga 3.

```
INSERT INTO [FTUNNUSED] ([arvutada],[FID],
[nimi],[ftyyp],[sumsimmax],[andmepäring],[vmaht],[kvmah], [tmaht])
VALUES (True,6,'enterobiasis',3,3,'Enterobiasis',200,500,6);
```

 **79. Lisa seletavate tunnuste 60...66 määratlused näidisandmete andmebaasi tabelist Lisatunnused tabelisse ATUNNUSED Constud andmebaasis.**

Accessis piisab ridade kopeerimisest. Lisaks tunnuse numbrile AID on ainus oluline väli nendes ridades KID, kuna tunnuste väärtused on tabelis juba olemas.

 **80. Lisa tabelisse ATUNNUSED väljad F6 ja AF6, mis vastavad muutujale FID = 6.**

Võid kasutada nuppu Väljad E... ja AF... tabelite loomise aknas või kopeerida ja muuta olemasolevaid välju või siis käivitada alltoodud päringu.

```
ALTER TABLE [ATUNNUSED] ADD [F6] bit NULL, [AF6] byte NULL;
```

 **81. Lülita [ATUNNUSED].[F6] = True ridades, kus AID > 59 ja [ATUNNUSED].[F6] = False, kus AID < 67.**

82. Lisa funktsioontunnuse logitabel LOG_F6.



Constud tabelite loomise akna nupp *Logitabelid* loob logitabelid kõigi muutujate jaoks, mis on kirjas tabelis *FTUNNUSED*.

83. Eemalda linnuke Constud põhiakna õppe osa lahtrist Ruumiandmed, säilita vaikumisi suvandid (joonis 4-3) ja käivita Constud õpe.



Kui soovid tulemusi kiiresti saada, siis vähenda õppekorduste arvu vaikumisi väärtusest (1000) väiksemaks. Umbes 200–500 õppekordust oleks järgmise ülesande jaoks siiski tarvis läbi teha. Veendu, et logitabelis *LOG_F6* ilmuvad esimeste õppekorduste tulemused ja jäta siis arvuti omapäi õppima. Õppe tulemused salvestuvad nagu binaarse muutuja puhulgi (peatükk 4.4).

4.9. MITMEMÕÕTMELINE MUUTUJA

Mitmemõõtmelise numbrilise muutuja hinnangute ja vaatluste vastavust mõõdetakse Constud süsteemis kõigi osamuutujate (dimensioonide) keskmise suhtelise ühisosa suurusega. Mitmemõõtmelise muutuja kasutus süsteemis Constud on alles arendamisel.

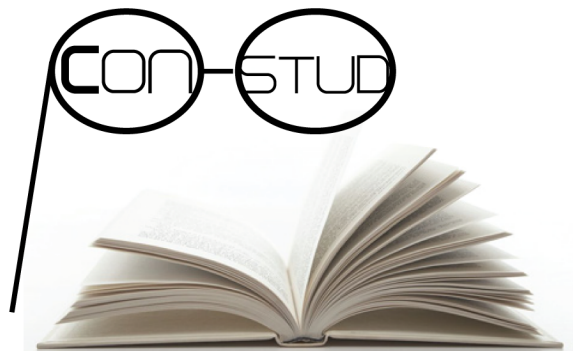


KÜSIMUSED

1. Mis on parim võimalik **vastavus** süsteemis **Constud** kaheväärtuselise muutuja puhul ja numbrilise muutuja puhul? Parim ei tähenda suurimat.
2. Millist tüüpi **sõltuval muutujal** on tavapäraselt korduvad väärtused **Constud** andmebaasi väljal **[FTUNNUSED].[FID]**?
3. Miks on vastavus õpetusvalimis aina parem kui kontrollvalimis? Kas teistpidi tulemus on ka võimalik?
4. Kas seni parimasse prognoosivasse komplekti kuuluvate tunnuste aktuaalsused peaksid alati suurimad olema?
5. Tehisintellekti süsteemidele on omane mälu ja kogemusest õppimine. **Constud** salvestab õppe käigus kaale ja aktuaalsusi. Mis mõttes võib aktuaalsusi nimetada **Constud** süsteemi mäluks, kaale aga mitte?
6. Millise **kontrollvastavuseni** jõudis soo-neiuvaiba esinemise ja puudumise õpe (peatükk 4.4), taimkatteüksuste kui **multinomiaalse muutuja** õpe (peatükk 4.5), sademete hulga kui numbrilise muutuja õpe (peatükk 4.6)?
7. Taimkatteüksuste eraldi õppes ei õnnestunud prognoosida kahe **vaatluse** klassikuuluvust — VID = 173 ja VID = 205. Põhjus tuleneb tunnuse AID = 13 väärtusest. Mis see põhjus on?
8. Millisesse ilmavaatlusjaama on **Constud** õppe tulemuste järgi kõige vähem sademeid prognoositud ja millises jaamas on tegelikult kõige vähem sademeid? **Constud** õppe tulemused varieeruvad, aga enamasti ei ole prognoositud ja tegelik miinimum samas jaamas. Miks?
9. Koosta enterobiaasi esinemise õppe tulemuste järgi seni parimat kontrollvastavust näitav õppeköver. Näidis on joonisel 4-6.

КОРОМАНЕ 5

КОРОМАНЕ



Teadmiste kontrolli osas on vaid kolm valikuvõimalust: **kontrollvastavuse** arvutus, **näidiste** uuestivalimine ja **logitabeli** sirvimine (joonis 5-1). Logitabeli sirvimine on abivahend logitabelis olevate õppetulemuste lehitsemiseks ja logitabelis lausena kirjas olevate tunnusekaalude salvestamiseks tabelisse, kus iga **seletava tunnuse** jaoks on omaette rida või veerg. Kontrollvastavuse arvutus on vajalik, kui õppes kasutatud kontrollvalim on väiksem kui **õpetusandmete** maht ning siis kui **vaatluste** hulk õpetusandmetes on pärast õpet muutunud. Näidiste uuesti valimisel kontrollib Constud ükshaaval iga vaatluse vajalikkust näidisenäidise ja temale sobivaimat kaalu. Näidiste uuesti valimise tulemusena näidiste komplekt muutub.

Indeksite arvutus	Tehisõpe	Teadmiste kontroll
<input checked="" type="checkbox"/> Ruumiandmed	<input checked="" type="checkbox"/> Kontrollvastavuse arvutus	
	<input type="checkbox"/> Näidiste uuestivalimine	
	<input type="checkbox"/> Logitabeli sirvimine	

Joonis. 5-1. Constud põhiakna teadmiste kontrolli osa.

Lülita ruumiandmed välja, kui vaatlustel puuduvad koordinaadid

5.1. LOGITABELITE SIRVIMINE

Logitabelite sirvimise vahend on loodud, et muuta Constud õppe tulemuste sirvimist mugavamaks ja võimaldada tunnusekaalude ümbersalvestamist andmebaasitabelist tekstifailis olevasse sõredasse tabelisse, kus iga tunnuse jaoks on kas omaette rida või omaette veerg.

84. Käivita Constud, vali andmebaas, millega töötasid, vali teadmiste kontrolli osa ja sealt Logitabeli sirvimine (joonis 5-1). Vajuta nuppu Jätka.



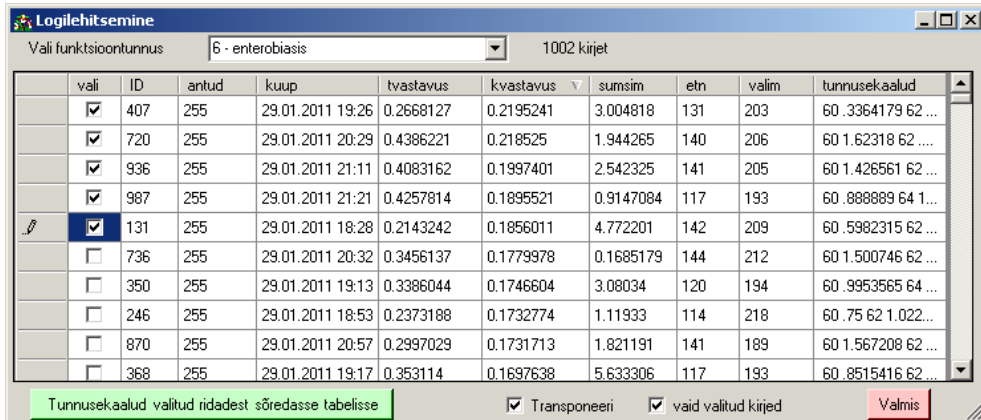
Avaneb logitabeli sirvimise aken (joonis 5-2).



85. Vali logitabeli sirvimise akna ülaosast funktsioontunnuseks 6 — enterobiaas (joonis 5-2).

Funktsioontunnuste nimed ja FID numbrid loetakse Constud andmebaasi tabelist **FTUNNUSED**. Kirjete arv ja lehitseja tabeli lahtrite sisu loetakse ülalt valitud funktsioontunnuse **logitabelist**. Lehitseja ridu saab ühes või teises veerus olevate väärtuste järgi järjestada ning tunnusekaalude veergu ja kogu akent saab laiemaks venitada.

Tunnuste kaalud on logitabelis kujul: <tunnuse AID number> <selle tunnuse kaal> <tunnuse AID number> <selle tunnuse kaal> jne.



vali	ID	antud	kuup	tvastavus	kvastavus	sumsim	etn	valim	tunnusekaalud
<input checked="" type="checkbox"/>	407	255	29.01.2011 19:26	0.2668127	0.2195241	3.004818	131	203	60.3364179 62 ...
<input checked="" type="checkbox"/>	720	255	29.01.2011 20:29	0.4386221	0.218525	1.944265	140	206	60.1.62318 62
<input checked="" type="checkbox"/>	936	255	29.01.2011 21:11	0.4083162	0.1997401	2.542325	141	205	60.1.426561 62 ...
<input checked="" type="checkbox"/>	987	255	29.01.2011 21:21	0.4257814	0.1895521	0.9147084	117	193	60.888889 64 1...
<input checked="" type="checkbox"/>	131	255	29.01.2011 18:28	0.2143242	0.1856011	4.772201	142	209	60.5982315 62 ...
<input type="checkbox"/>	736	255	29.01.2011 20:32	0.3456137	0.1779978	0.1685179	144	212	60.1.500746 62 ...
<input type="checkbox"/>	350	255	29.01.2011 19:13	0.3386044	0.1746604	3.08034	120	194	60.9953565 64 ...
<input type="checkbox"/>	246	255	29.01.2011 18:53	0.2373188	0.1732774	1.11933	114	218	60.75 62 1.022...
<input type="checkbox"/>	870	255	29.01.2011 20:57	0.2997029	0.1731713	1.821191	141	189	60.1.567208 62 ...
<input type="checkbox"/>	368	255	29.01.2011 19:17	0.353114	0.1697638	5.633306	117	193	60.8515416 62 ...

Joonis. 5-2. Constud logitabeli sirvimise aken. Esimeses veerus on valitud viis kontrollvastavuse järgi paremat tulemust.

86. Järjesta kirjed kontrollvastavuse järgi ja vali esimesse veergu linnukest klõpsates viis kontrollvastavuse järgi parimat rida (joonis 5-2).

87. Proovi valitud viie parema õppekorduse tunnusekaale nii sõreda tabeli ridadesse (transponeeri) kui ka veergudesse teisendada.

Tekstifaili eksporditud tunnusekaalude ja õppekorduste risttabeli saab avada Excelis, kus on lihtne ridu- ja veergepidi kokkuvõtteid arvutada. Pane tähele, et kasutatud tunnuste järjekord keskmise kaalu järgi viies parimas õppekorduses ei ole sama, mis tunnuste järjekord aktuaalsuste järgi ning samuti, et numbrilise muutuja puhul näitab väiksem vastavusearv väiksemat ruutkeskmist viga, teiste muutujate puhul on paremad need vastavused, mis on arvuliselt suuremad (ühele lähemal).

5.2. KONTROLLVASTAVUSE ARVUTUS

Valimite kasutamisel näitab parim õppes saadud vastavus ühes õppevalimis saadud tulemust ning prognoositud väärtused ja nende järgi arvutatud kontrollvastavus ühes kontrollvalimis arvutatud vastavust. Ülejäänud vaatlustele salvestatud hinnangud on saadud õppe käigus mõnest eelnevast tookord parimast kontrollvalimist. Selleks, et saada parima komplekti järgi prognoositud väärtused kõigile vaatlustele ja et saada kogu andmestikus kehtiv parima prognoosiva komplekti kontrollvastavus, tuleb käivitada kontrollvastavuse arvutus (joonis 5-1).

Kui kontrollvalim < õpetusandmed, siis pärinevad Constud õppes salvestatud prognoositud väärtused erinevatest kontrollvalimitest

Kui Constud andmebaasi selles vaatluste tabelis, mis on seotud parajasti kontrollitava funktsioontunnusega, on olemas väli [exemplars_used], siis salvestatakse sinna iga vaatluse puhul selle vaatluse prognoosimisel kasutatud näidiste VID numbrid ja sarnasused. Kasutatud näidiste järgi saab teada, kuidas üks või teine prognoositud väärtus tekkis. Kasutatud näidiste uurimine võib anda selgust, miks mõnedes vaatluskohtades on prognoositud väärtused üllatavalt väärad. Võib selguda, et need väärad otsused on Constud täiesti loogiliselt tuletanud ühes või teises mõttes ebaesinduslike või vigaste vaatluste järgi, mis on näidiseks valitud. Tehisõppesüsteem ei saa kuidagi vastutada õpetusandmete esinduslikkuse eest ega välikaardistuse õigsuse eest.

Constud ei vastuta õpetusandmete õigsuse ega esinduslikkuse eest



88. Tee koopia Balti sademete tabelist ja logitabeli parima kontrollvastavuse reast.

Balti sademete logitabeli nimi sisaldab sama FID numbrit, mis on Balti sademetel funktsioontunnuste tabelis (joonis 4-5). Võid kopeerida ka kogu logitabeli. Koopiade loomine on vajalik, sest kontrollvastavuse arvutamise lõpus Constud uuendab seni parima õppekorduse tulemused. Kas samasse andmebaasi või kuhugi mujale kopeeritud tabelitest saab vaadata õppe tulemusi, mis olid enne kontrollvastavuse arvutust.

89. Vali funktsioontunnuste tabelis väljalt [arvutada] Baltimaade sademete uurimise rida ja lülita välja kõik teised read.



90. Käivita kontrollvastavuse arvutus ja võrdle seda esialgse vastavusega logitabelist kopeeritud real. Võrdle ka tunnuste kaale logitabelis.

Kuna sademete hulga õppes olid kõik 123 vaatlusjaama kontrollvalimisse kaasatud, siis uuenes vaid parima vastavuse salvestamise aeg logitabelis.




91. Võrdle näidiste kaale ja prognoositud väärtusi õpetusandmete tabeli kahes versioonis.

Selleks on mugav kasutada päringut, kus võrreldavad veerud kuvatakse kõrvuti,

nagu näiteks järgmine päring. Jälgi, et kopeeritud tabeli nimi oleks õige.

```
SELECT [Baltisademed].[progn],  
[Koopia objektist Baltisademed].[progn],  
[Baltisademed].[w], [Koopia objektist Baltisademed].[w]  
FROM [Koopia objektist Baltisademed]  
INNER JOIN [Baltisademed] ON  
[Koopia objektist Baltisademed].[VID] = [Baltisademed].[VID];
```


Kui Constud õppe kontrollvalim hõlmab kõiki
õpetusvaatlusi, siis kontrollvastavuse eraldi arvutamine
tulemusi ei muuda

92. Lisa tabelisse *Baltisademed* viie vaatlusjaama andmed andmebaasi *Constud_õppe_andmed.mdb* tabelist *Baltisademed_pluss*. 

Selle muudatusega imiteerime olukorda, kus tehiseõppesüsteem on mingitele õpetusandmetele sobitatud, siis aga lisandub õpetusandmeid juurde ja oleks tarvis teada, kui palju tuleks hinnangute oodatavat täpsust lisandunud vaatluste tõttu korrigeerida.


93. Käivita kontrollvastavuse arvutus ja võrdle seda esialgse kontrollvastavusega.  
Võrdle ka tunnuste kaale logitabelis.

Seekord peaks parima tulemuse **kontrollvastavus** muutuma.

94. Võrdle näidiste kaale ja prognoositud väärtusi õpetusandmete tabeli kahes versioonis. 

Kontrollvastavuse eraldi arvutamine ei muuda näidiste ega tunnuste kaale. Kuna parim prognoosiv komplekt ei muutunud, siis ei saa muutuda ka prognoositud väärtused.

Kontrollvastavuse arvutamine ei muuda näidiste ega
tunnuste kaale

95. Lisa Balti sademete õpetusandmete tabelisse ja päringusse teksti tüüpi väli nimega *Exemplars_used*. 

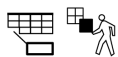
Välja suuruseks võiks määrata 255 tähemärki.

96. Käivita kontrollvastavuse arvutus veelkord. 

Väljale *Exemplars_used* peaks iga vaatlusejaama kohas salvestuma nende näidiste VID numbrid, mille järgi seda jaama prognoositi ja iga näidise sarnasus prognoositava kohaga. Numbrilise muutuja prognoositud väärtus arvatati näidistel olevate väärtuste kaalutud keskmisena, kusjuures kaaludeks olid sarnasused.

Taimkatteüksuste õpe erines sademete hulga õppest valimite kasutamise poolest. Veendu järgmistes ülesannetes **funktsioontunnusele** FID = 3 kontrollvastavust arvutades, et prognoositud taimkatteklassid **õpetusandmete** päringus muutuvad. Õppe ajal arvatati kontrollvastavus kontrollvalimis, kontrollvastavuse arvutuse eraldi käivitamisel kaasatakse kõik **vaatlused**. **Eelklassifikaatori** kasutamisel jääb hinnang arvutamata vaid nendele vaatlustele, mis on ainsad oma eelklassis.

Ühtegi vaatlustulemust ei saa prognoosida tema enda järgi



97. Loo koopia taimkatteüksuste tabelist *Taimkate* ja tabeli *LOGI_F3* reast, mille number on tabelis *FTUNNUSED*. Lisa tekstitüüpi väli *[Taimkate].[Exemplars_used]*.

Võid kopeerida ka kogu tabeli *LOGI_F3*.

98. Lülita **funktsioontunnuste tabelis sisse vaid muutuja 3**.

99. Käivita **kontrollvastavuse arvutus**.

100. Võrdle **näidiste kaale ja prognoositud väärtusi taimkatteüksuste tabeli kahes versioonis**.

Prognoositud väärtused muutusid ja need salvestati kõigile **vaatlustele**. Kui mõnele vaatlusele ei saanud hinnangut arvutada, siis salvestati määramatuse kood — 255. Määramatus tekib siis, kui ühtegi sarnast **näidist** ei ole. **Eelklassifikaatori** kasutamisel on määramatuse sage põhjus näidiste puudumine selles eelklassis. Prognoositud väärtused muutuvad nendel vaatlustel, mis ei olnud parima **õppekorduse** kontrollvalimis — nende hinnangud pärinesid mõnest eelmisest õppekordusest.

Eelklassifikaatori kasutamisel peavad igas eelklassis olema omad näidised

5.3. NÄIDISTE UUESTI VALIK

Kontrollvastavuse arvutus võimaldab prognooside oodatavat täpsust andmete lisandumise või muul moel muutumise järel ümber hinnata. Kui on põhjust arvata, et õpetusandmete uuenemine võimaldab leida senisest paremat tunnuste ja näidiste kaalude komplekti, siis tuleb tehisõpe uuesti käivitada. Constud praegune versioon pakub vaid näidiste uuesti valimise võimalust, mitte tunnuste uuesti valimist. Tunnuste muutumisel tuleks taaskäivitada Constud õpe.

101. Tee veel üks koopia Balti sademete tabelist ja logitabeli parima vastavuse reast või kogu logitabelist.



102. Lülita funktsioontunnuste tabelis sisse vaid Balti sademete muutuja. Käivita näidiste uuestivalimine.



103. Võrdle ka näidiste ja tunnuste kaale ning õpetusandmetele prognoositud väärtusi enne ja pärast näidiste uuestivalikut.



Näidiste komplekt ja prognoositud väärtused muutusid täielikult. Prognoositäpsus halvenes. Näidiste uuesti valiku algoritm vajab parendamist.

WINNANGUD

6

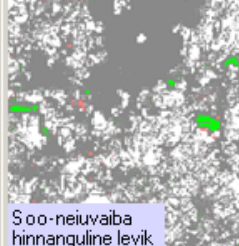
108	5	1	42	2		58	160	0	77	4	1	7	33		79	6		6
5	2	44	3	##	#	55	45	##	52		##	8		##	2	5	7	6
44	2	88	###	6	68	6	##	4	##	44	##	##		##	4	55	9	
2		44	###	7	99	78	##	48	##	0	##	##		##		5		33
45	48	3	###	6		8	##		##	45	##	4	##	##	333		6	5
4	45	99		##	#	46		##		76	##	8		##	6	6	42	33
15	8	1	4	1	0	79		0	0		7	0		166	0	77		7
48		4	4	0	4	0	11	0	46	0	0	55	66	88	4	150	79	44
	19	66	###	99	#	###	###	###	#	##	5	##	5	##	##		79	
46	13	###	2	18	21	0	###	2		##	2	##	5	##		#	89	7
49		###	8	0	8	0	###	85	1	##	89	##	4	##	11	##	99	44
2	3	78	###	0,5	45	15	###	22	0	##	65	##		##	55	##		2
15		5	0	###	413		###	2		##	0	##	88	##	11	##	2	44
45	13	1		###	22		###	89	0	##	2	##	0	##		#	99	1
45		65	###	22	60	120	###	2		#	###	#	0	##	##		0	1
22	22		7	22	12	7	22		44			22		0	25		22	4
0	7	545	2	22	21	8	2	7	22	1	212		212	0	0	89		1

Hinnangute tabelisse arvutamisel tuleb Constud põhiakna hinnangute arvutamise osas valida: 1) kas **vaatlused** on seotud asukoha koordinaatidega, 2) kas arvutatakse hinnang ja sarnasus **näidistega** või 3) arvutatakse sarnasust etteantud kategooriaga (joonis 6-1). Vaatluste asukohakoordinaate prognoosil ja sarnasuste arvutamisel ei kasutata, kuid alles jääb erinevus Constud andmestruktuuris: ruumandmete puhul määrab **argumenttunnuse** tüüpi (nominaalne/numbriline) indeks, mitteruumiliste andmete puhul andmekiht, mitteruumilistele vaatlustele ei rakendata ajalise kehtivuse piiranguid.

Nullsarnasuse tase ja sarnasuse ulatus **standardhälvetes** peaksid prognooside arvutamisel olema samad, mis olid õppes (peatükk 4.1.1 ja 4.2). Vaikimisi väärtused null ja kaks peaksid enamiku andmete jaoks kõige paremini sobima.

Hinnangute arvutamisel andmebaasi on oluline tähele panna, et **tehisõppes** valitud näidised, nende tunnused ja kaalud asuvad **andmebaasiobjektis**, mille nimi on Constud andmebaasi väljal **[FTUNNUSED].[andmepäring]**; prognoositavad vaatlused on aga teises andmebaasiobjektis ja selle nimi tuleb kirjutada hinnangute arvutamise aknas olevasse lahtrisse (joonis 6-1).

Prognoosi arvutamiseks tuleb **funktsioontunnuse** tabelis sisse lülitada prognoositava muutuja rida. Seesama väli, mida kasutati selle muutuja õppes. Prognoositavatel vaatlustel peavad olema väljad vähemalt nendele **seletavatele tunnustele**, mis on parimasse prognoosivasse komplekti valitud. Muidugi ka vaatluse number — **VID**, väli **progn** prognoositud väärtuste salvestamiseks ja väli **prob** näidistega sarnasuse salvestamiseks. Sarnasuse arvutamisel etteantud kategooriaga peab hinnatavate vaatluste tabelis olema väli **F**, mis sisaldab selle kategooria koodi, millega sarnasust hinnatakse. Prognoosimiseks vajalikud tunnusekaalud loetakse **logitabeli** realt, mille number on väljal **[FTUNNUSED].[LogiID]**.

Indeksite arvutus	Tehisõpe	Teadmiste kontroll	Hinnangute arvutus
 <p>Soo-neiuvaiba hinnanguline levik</p>	<input checked="" type="checkbox"/> Ruumiandmed <input type="checkbox"/> SD tabelist STDEV Nullsarnasus: <input type="text" value="0"/> Sarnasuse ulatus [SD]: <input type="text" value="2"/>	<input checked="" type="checkbox"/> Ruut-kaardilehed piksli külj: <input type="text" value="10000"/>	Constud andmebaasi <input checked="" type="checkbox"/> Hinnang tabelisse: <input type="text" value="HINNANGUD"/> <input checked="" type="checkbox"/> Sarnasus etteantud üksusega tabelisse: <input type="text" value="SARNASUSTABEL"/>
	Rasterfaili		<input type="checkbox"/> Üksikmuutuja hinnang <input type="checkbox"/> Komplekshinnang

Joonis 6-1. Hinnangute arvutamise **suvid** Constud põhiaknas.

6.1. KAHEVÄÄRTUSELINE MUUTUJA

Eelmistes peatükkides on ette valmistatud kahe erineva kaheväärtuselise muutuja

andmestikud: soo-neiuvaiba esinemine/puudumine ja enterobiaasi esinemine/puudumine. Neist esimese andmestiku **vaatlused** on asukohakoordinaatidega seotud ja vaatlusi kirjeldavad kohatunnused. Teine andmestik on mitteruumilise andmestiku näidis.

6.1.1. RUUMILINE KAHEVÄÄRTUSELINE MUUTUJA

Kohad, mille puhul oleks tarvis hinnata, kas need sarnanevad rohkem soo-neiuvaiba esinemiskohtadega või puudumiskohtadega, on näidisandmebaasi tabelis *Prognoositavad_kohad*.

104. Impordi tabel *Prognoositavad_kohad* Constud andmebaasi.



Pane tähele, et imporditud tabelis on vaid kirjete numbrid ja koordinaadid. Vajalikud väljad tuleb seekord ise lisada ja kohta kirjeldavate tunnuste väärtused arvutada.

105. Lisa tabelisse *Prognoositavad_kohad* ruumilise muutuja prognoosi arvutamiseks vajalikud väljad: [progn], [prob], [alates], [kuni] ning selle muutuja seletavate tunnuste väljad. Seejärel lülita väljal [FTUNNUSED].[arvutada] sisse sõltuv muutuja E. palustris on/pole ja kõik teised muutujad välja.



Soo-neiuvaiba esinemise/puudumise õppes kasutasime tunnuseid AID = 1...9 ja AID = 11...12. Prognoosi arvutamiseks on vajalikud vaid need tunnused, mis on õppe tulemusel selgunud parimas prognoosivas komplektis veerus *tunnusekaalud logitabeli* real, mille ID number on tabelis *FTUNNUSED*. SQL kujul oleks väljade lisamise päring umbes järgmine. Kuna Constud õppe tulemused varieeruvad, siis on kindlam lisada kõik õppes osalenud tunnused.

```
ALTER TABLE [Prognoositavad_kohad] ADD
[progn] byte NULL, [prob] byte NULL, [alates] byte NULL, [kuni] byte
NULL, [1] byte NULL, [2] byte NULL, [3] byte NULL, [4] byte NULL, [5]
byte NULL, [6] byte NULL, [7] byte NULL, [8] byte NULL, [9] byte
NULL, [11] byte NULL, [12] byte NULL;
```

```
UPDATE [FTUNNUSED] SET [arvutada]= False;
```



```
UPDATE [FTUNNUSED] SET [arvutada]= True WHERE FID=2;
```

106. Kontrolli, et väljal [ATUNNUSED].[F2] oleksid sisse lülitatud vaid tunnused AID = 1...9 ja AID = 11...12 ning seejärel vali Constud indeksite arvutus prognoositabelisse. Tabeli nime lahtrisse märgi *Prognoositavad_kohad* või muudetud nimi, kui oled prognoositavate vaatluste tabeli nime muutnud (joonis 6-2). Käivita arvutus.





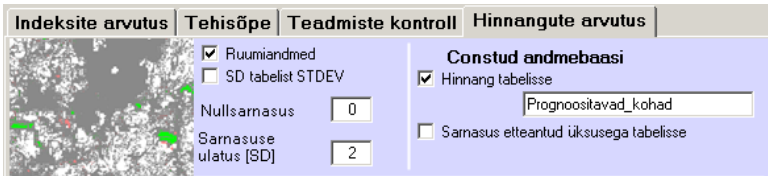
Joonis 6-2. Constud **su**vandid indeksite arvutamisel prognoositavate vaatluste tabelisse.

  **107. Veendu, et seletavate tunnuste väärtused jõudsid prognoositavate vaatluskohtade tabelisse ja käivita Constud hinnangute arvutus (joonis 6-3).**

Arv üks väljal *progn* näitab koha suuremat sarnasust liigi esinemise näidiskohtadega, arv null näitab, et koht sarnaneb rohkem liigi puudumise näidiskohtadega. Väljal *prob* on koha keskmine protsentuaalne sarnasus nende näidistega, mille järgi hinnang saadi. Sarnasust hinnati nende tunnuste järgi, mis on kirjas **logitabelis** parima **kontrollvastavusega** õppevooru real. Kui sarnasus = 100, siis on koht kasutatud tunnuste poolest näidistega identne.

Sarnasus näidistega on otsusekindlus, mitte sarnasus esinemiskohtadega

Näidiste kirjeldus ja kaalud on **õpetusandmete** tabelis, mille nimi on väljal **[FTUNNUSED].[andmepäring]**.



Joonis 6-3. Constud **su**vandid hinnangute arvutamisel tabelisse.


6.1.2. MITTERUUMILINE MUUTUJA


Mitte-ruumilise muutuja näitena õppis Constud peatükis 4.8 eristama enterobiaasihaigeid ja terveid lapsi Kagu-Eesti lasteaeadeadest. Õpitu rakendamist võiks proovida kolme lapse näitel, kelle teadaolevad tunnused on tabelis 1. Nende laste kohta hinnangute arvutamiseks süsteemis Constud on tarvis luua **andmebaasiobjekt** peatüki 6 alguses loetletud väljadega ning siis tabelis olevad väärtused sinna sisse kanda. Üks lihtne moodus prognoositavate vaatluste tabeli loomiseks on teha koopia sama muutuja õpetusandmete tabeli struktuurist, ilma andmeteta. Väljad *F*, *w* ja *aktuaalsus* võib kustutada, need ei ole hinnangu arvutamisel vajalikud.


Tabel 1. Laste andmed enterobiaasi nakatatusse sarnasuse järgi hindamiseks.



AID	60	61	62	63	64	65	66
VID	Vanus	Sugu	Laste arv peres	Ühevanuselise rühm	Vanusehaare rühmas	Piirkond	Asula tüüp
1100	2	0	1	Jah	1	1	1
1101	4	1	3	Ei	3	3	3
1102	6	1	5	Ei	4	5	5



Sugu: 0 — tüdruk, 1 — poiss; piirkond: 1 — Tartu linn, 3 — Tartu maa-asula, 5 — Põlva maa-asula.

108. Kopeeri tabeli *Enterobiasis* struktuur tabeliks kolmlast (muidugi võid uuele tabelile ka mingi muu nime anda). Kustuta tabelist kolmlast väljad F, w ja aktuaalsus. 

109. Kanna selle juhendi tabelis 1 olevad andmed Constud andmebaasi tabelisse kolmlast. 

110. Lülita väljal [FTUNNUSED].[arvutada] sisse sõltuv muutuja enterobiaas ja kõik teised muutujad välja. 

111. Ava Constud. Vali hinnangute arvutus. Lülita välja ruumiandmed. Kirjuta hinnangu tabeli lahtrisse selle tabeli nimi, kus on prognoositavad vaatlused (kolmlast).  

112. Käivita hinnangute arvutus ja seejärel vaata prognoositud väärtuste tabeli (kolmlast) väljalt progn, millised kolmest lapsest sarnanevad enterobiaasahaigetega ja millised tervetega.  

Haiged on väljal *progn* kodeeritud väärtusega *Jah*, haiged *Ei*. Sarnasus väljal *prob* näitab keskmist sarnasust prognoositud klassi näidistega.

6.2. MULTINOMIAALNE MUUTUJA

Multinomiaalse ja pideva muutuja prognoos ei erine põhimõtteliselt eelmistest näidetest. Endiselt peab olema eelnevalt läbitud tunnuse õpe, mille käigus on leitud parim prognoosiv komplekt. Funktsioontunnuste tabelis tuleb sisse lülitada prognoositava muutuja arvutamine, käivitada Constud, valida ruumiandmed või need välja lülitada ja sisestada prognoositavate vaatluste tabeli, päringu või vaate nimi. Proovigem seda näidisandmetega. Taimkatteüksuste õpetusandmeid piirasime

tingimusega VID < 301. Tabelis *Taimkate* on aga 500 vaatluskoha kirjeldavad tunnused ja nendes kohtades välivaatlusel määratud taimkatteüksuste koodid. See annab võimaluse võrrelda sarnasuse järgi tuvastatud taimkatteklasse välivaatluste tulemustega.


- 113. Lülita väljal [FTUNNUSED].[arvutada] sisse sõltuv muutuja FID = 3 (taimkate) ja kõik teised muutujad välja.



- 114. Loo andmebaasivaade, mis valiks taimkattevaatlused VID numbriga > 300. Salvesta see vaade ja anna sellele nimi, näiteks *Taimkatte_test*.

```
SELECT * FROM [Taimkate] WHERE VID>300;
```

- ↻ 115. Ava Constud. Vali hinnangute arvutus. Jälgi, et ruumiandmed oleksid sisse lülitatud. Kirjuta hinnangutabeli lahtrisse selle tabeli nimi, kus on prognoositavad vaatlused (*Taimkatte_test*).

- ↻  116. Käivita hinnangute arvutus ja seejärel vaata prognoositud kohtade tabeli veergudest F, progn ja prob, kui sageli langes sarnasuse järgi hinnang kokku välivaatlusega ja kui suur on sarnasus näidistega.

Kokkulangevust on üsna harva. Kui kasutada üle saja taimkatteüksuse, siis on keeruka loodusmaastiku enamikus kohtades isegi välivaatlejate vahel eriarvamused. Eristada olemasolevate kaardi- ja kaugseireandmete järgi nii suurt hulka üksusi kaardi kasutajat rahuldava tõepärasusega on üsna lootusetu ettevõtmine.

6.3. PIDEV MUUTUJA

Oluline erinevus teistest muutujatüüpidest on, et pideva muutuja puhul peab prognooside salvestamise väli *progn* olema neljabaidise murdarvu vormingus. Kui hinnangut arvutada ei saa, salvestab Constud prognoositud väärtuseks muutuja keskmise õpetusandmetes. Teiste muutuja tüüpide puhul salvestatakse määramatuse korral arv 255.

Näidisandmete andmebaasis on tabel *Baltisademed_test*, kus on keskmine aastane sademete hulk millimeetrites 117 vaatlusjaamas, jaama asukohta kirjeldavad tunnused, mida kasutasime ka sademete hulga õppes ning väljad prognoositud väärtustele.

- 117. Lülita väljal [FTUNNUSED].[arvutada] sisse sõltuv muutuja FID = 4 (Balti sademed) ja kõik teised muutujad välja.



- 118. Impordi või kopeeri näidisandmete andmebaasist Constud andmebaasi tabel

Baltisademed_test.

119. Ava Constud. Vali hinnangute arvutus. Jälgi, et ruumiandmed oleksid sisse lülitatud. Kirjuta hinnangu tabeli lahtrisse selle tabeli nimi, kus on prognoositavad vaatlused (Baltisademed_test).



120. Käivita hinnangute arvutus ja seejärel vaata prognoositud kohtade tabelist, kui suur on hinnangu ja vaatluste erinevus. Kas prognoositud sademete hulk on suurem samades jaamades, kus vaadeldud sademete hulk on suurem?



6.4. MULTINOMIAALSE MUUTUJA ERALDI ÕPITUD ÜKSUSED

Multinomiaalse muutuja üksuste ükshaaval õppe jaoks loodi harjutuses 68 andmebaasivaade *Taimkatteõpe1*, mis loeb 300 esimest vaatlust tabelist *Taimkate* ja nimetab ümber väljad: *progn1*, *prob1*, *aktuaalsus1*, *w1*, andes neile Constud eestikeelse andmebaasi struktuurile vastavad nimed: *progn*, *prob*, *aktuaalsus*, *w*.

121. Kopeeri andmebaasivaade *Taimkatteõpe1*, anna koopiale nimi (näiteks *Taimkatte_test1*) ja muuda seda nii, et valitaks kirjed, mille VID > 300.



```
SELECT [VID],[F],[X],[Y],[progn1] AS progn, [prob1] AS prob,
[Alates], [Kuni], [13],[14],[15],[16],[17],[18],[19],[20],[21],[22],
[23], [24],[25],[26],[27],[28],[29],[30],[31],[32],[33],[34],[35],
[36], [37],[38],[39],[40],[41],[42],[43],[44],[45],[46],[47],[48],
[49],[50]
FROM [Taimkate] WHERE [VID]>300;
```

Väljad *w* ja *aktuaalsus* võib ära jätta. Väli *F* on praegusel juhul vajalik vaid prognoositud üksuste võrdlemiseks vaatlustulemustega. Constud prognoosi arvutamisel välja *F* ei kasuta. Pane tähele, et päring nimetab ümber väljad *progn1* ja *prob1* ning edastab vaid kirjed, millel VID > 300. Esimest 300 kirjet kasutati *õpetusandmetena*.

122. Lülitä väljal [FTUNNUSED].[arvutada] sisse kõik read, kus FID = 5 ja kõik teised muutujad välja. Kontrolli, et tunnus 13 oleks märgitud eelklassifikaatoriks ridades, kus FID = 5.



123. Ava Constud. Vali hinnangute arvutus. Jälgi, et ruumiandmed oleksid sisse lülitatud. Kirjuta hinnangu tabeli lahtrisse selle tabeli nimi, kus on prognoositavad vaatlused (*Taimkatte_test1*).



124. Käivita hinnangute arvutus ja seejärel tutvü prognoositud tulemustega.



Pane tähele, et määramatuse kood 255 on kirjetel, millel on eelklassifitseeriva tunnuse 13 (põhikaardi põhiala) teatud väärtused. Nende koodidega põhialade kohta

ei ole prognoosivas komplektis ühtegi näidist ja ei ole millegi järgi taimkatteüksust tuvastada.

Prognoositud koodid esindavad vaid neid üksusi, mis on funktsioontunnuste tabelis valitud. Seekord oli tehisõppes ja prognoosis viis eraldi taimkatteüksust. Kõik vaatluskohad, millel oli vähimgi sarnasus nende viie üksuse vähemalt mõne näidisega, klassifitseeriti neist üksusest kõige sarnasemasse. Kui Constud prognoosi arvutuse käivitamisel on valitud vaid üks üksus, siis prognoositakse kõik kohad, mis vähegi selle üksuse näidistega sarnanevad, sellesse üksusesse. Seetõttu, kui lasta Constud süsteemil õppida igat üksust eraldi, siis on mõtet ka kõiki neid üksusi prognoosis koos kasutada.

6.5. SARNASUS ETTEANTUD ÜKSUSEGA

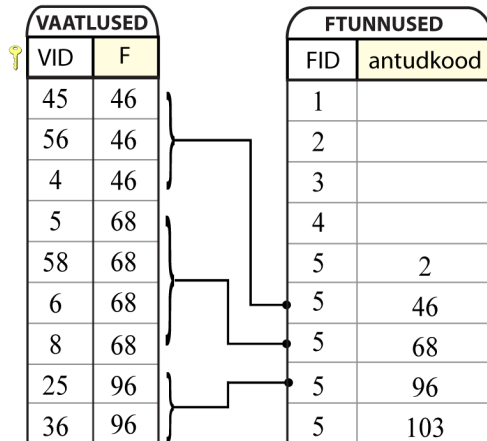
Sarnasus etteantud üksusega võimaldab hinnata, kui sarnane on üks või teine vaatlus etteantud üksuse näidistega. Kaardi kontrollimisel võimaldab see hinnata iga koha sarnasust kaardil oleva üksuse näidistega kaugseireandmetest arvatud tunnuste poolest. Kohas, kus sarnasus on väike, on kas kaardistus vale, koht on eriliselt omanäoline, maastik on muutunud või siis ei ole etteantud üksuse näidiste komplekt piisavalt esinduslik.

Sarnasuse hindamisel eeldatakse, et funktsioontunnuste tabelis on arvutamiseks valitud vaid sama FID numbriga read. Kui valitud on rohkem või vähem, siis annab Constud veateate ja ootab vea parandamist. Arvutada saab sarnasust vaid valitud muutuja nende üksustega, mille äratundmist on Constud eraldi õppinud ($ftyyp = 4$) ja mis on valitud väljal [FTUNNUSED].[arvutada]. Etteantud üksusega sarnasuse hindamisel kõige sarnasemat klassi ei otsita ja igal üksusel on oma prognoosiv tunnuste ja näidiste komplekt.


Tabelis, kus on vaatlused, mille sarnasust antud üksusega hinnatakse, peab lisaks hinnangu arvutamisel vajalikele väljadele olema väli F , millel on selle üksuse kood, millega sarnasust tuleb hinnata (joonis 6-4). Sarnasustabel erineb prognoositavate vaatluste tabelist vaid välja F vajalikkuse ja välja *progn* mittevajalikkuse poolest.

Constud eeldab, et kõigil üksustel on sama FID number, samad õpetusandmed ja eelklassifikaator, logitabel muidugi ka. Sarnasus hinnatakse vaid nende vaatluste puhul, millele omistatud antudkoodi väärtusele vastav rida on funktsioontunnuste tabelis sisse lülitatud; eelklassifikaatori kasutamisel arvutatakse sarnasus vaid nendes eelklassides, kus on selle kategooria näidiseid.

Järgnevates harjutustes taaskasutatakse juba loodud andmebaasiobjekti *Taimkatte_test1*.






Joonis 6-4. Seos välja *[FTUNNUSED].[antudkood]* ja välja *F* vahel. Väljal *F* on igale uuritavale vaatlusele omistatud etteantud kood.

125. Lülita väljal *[FTUNNUSED].[arvutada]* sisse rida, kus *FID = 5* ja uuritav üksus on sööt ning kõik teised read välja. 

126. Tühjenda sarnasuste väli andmebaasiobjektis *Taimkatte_test1*. 

```
UPDATE Taimkatte_test1 SET [prob] = Null;
```

127. Ava Constud. Vali sarnasuste arvutus etteantud üksustega. Jälgi, et ruumiandmed oleksid sisse lülitatud. Kirjuta sarnasustabeli lahtrisse tabeli nimi, kus on prognoositavad vaatlused (joonis 6-5). Käivita sarnasuse arvutus. Jäta Constud avatuks. Värskenda uuritavate vaatluste loetelu.  

128. Järjesta *Taimkatte_test1* välja *F* järgi ja liigu andmetabelis piirkonda, kus on söödi vaatlused (*F = 96*) või siis vali söödi vaatlused alltoodud päringuga. 

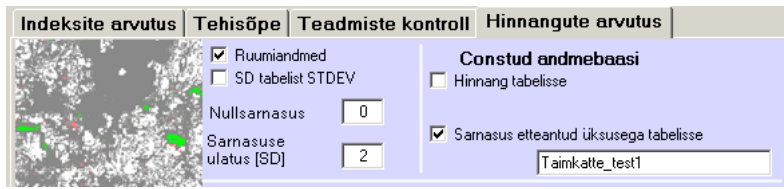
```
SELECT [F], [progn],[prob],[13] FROM [Taimkatte_test1] WHERE [F]=96;
```

Osades põhikaardi põhialades (tunnus 13) on sarnasused nullid, sest nende kohta ei ole prognoosivas komplektis ühtegi näidist ja põhikaardi põhialad olid eelklassifikaatoriks.

129. Eemalda eelklassifikaator tabeli *FTUNNUSED* ridadest, kus uuritava muutuja *FID = 5*. 

Constud praegune versioon ei võimalda sama funktsioontunnuse eraldi õpitud klassides erinevat eelklassifikaatorit kasutada.

```
UPDATE FTUNNUSED SET [eelklass]=Null WHERE FID=5;
```



Joonis 6-5. Constud **suvandid** etteantud üksusega sarnasuse arvutamisel tabelisse.

130. Käivita sarnasuse arvutus uuesti. Värskenda uuritavate vaatluste loetelu.

Nüüd ilmused sarnasused ka teiste tunnus 13 väärtuste puhul.

131. Lülita sisse kõik üksused FID = 5 puhul.

```
UPDATE FTUNNUSED SET [arvutada]=True WHERE FID=5;
```

132. Käivita sarnasuse arvutus uuesti. Jäta Constud avatuks. Värskenda uuritavate vaatluste loetelu.

Sarnasus arvutati kõigile vaatlustele, mille väljal *F* olev kood on esindatud tabeli *FTUNNUSED* väljal *antudkood*.

Järgmisena proovime arvutada kõigi vaatluste sarnasust ühe etteantud üksusega, näiteks söödiga.

133. Märki kõigi uuritavate vaatluste väljale F sama üksus, näiteks 96 (sööt).

```
UPDATE [Taimkatte_test1] SET [F]=96;
```

134. Käivita sarnasuse arvutus veelkord. Värskenda uuritavate vaatluste loetelu.

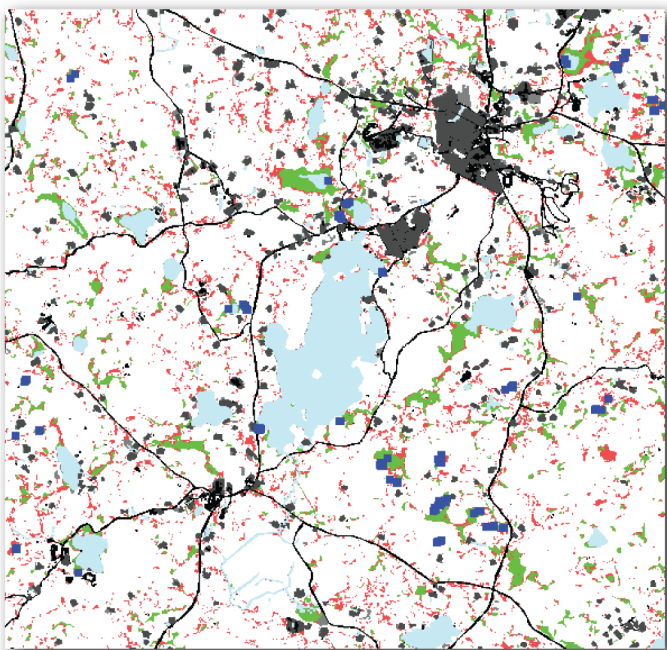
Tulemuste tabelisse ilmused iga koha sarnasus söödiga, arvestades söödi tuvastamiseks sobitatud tunnuste ja **näidiste** kaale ning eirates põhikaardi põhialasid (mida tegelikult arvestati prognoosiva komplekti õppes).



KÜSIMUSED

1. Millises vahemikus saab olla kaheväärtuselise prognoositava muutuja sarnasus näidistega?
2. Millised tabelis 1 loetletud lastest on tõenäoliselt enterobiaasi nakatunud, millised mitte?
3. Milliseid tunnuseid kasutades jagas Constud ülesandes 111 lapsed haigeteks ja terveteks? Märki need tunnused, mis osalesid prognoosis, mitte need, mis osalesid õppes.
4. Mitmel protsendil hinnatud vaatluskohtadest langes ülesandes 116 sarnasuse järgi hinnatud taimkatteklass kokku väливаatlusel määratud taimkatteklassiga?
5. Milline on keskmine sarnasus näidistega ülesandes 116 hinnatud vaatluskohtades, kus taimkatteüksust õnnestus ära tunda ja nendes kohtades, kus ei õnnestunud? Jäta määramatuse koodiga (255) kohad arvestusest välja.
6. Kui suur on aasta sademete hulga keskmise prognoositud väärtuse ja keskmise vaadeldud väärtuse erinevus (ülesanne 120)?
7. Kumb varieerub rohkem, kas vaadeldud sademete hulk või sarnasuse järgi prognoositud sademete hulk? Miks? Varieeruvuse mõõtmiseks sobib eelkõige haare ja **standardhälve**.
8. Millised põhikaardi põhiala vaatluskohad sarnanevad söödiga kõige enam ülesanne 134 tulemuste kohaselt? Põhialade koodid on näidisandmete andmebaasi tabelis *Põhialad*.
9. Mis mõttes moodustavad peatükis 6.2 prognoositavad vaatlused sõltumatu valimi?

КЛЯДЮ 7



Hinnanguliste kaartide loomiseks süsteemis Constud peab olema: 1) Constud andmebaas tunnuste ja näidiste kaalude komplektiga, 2) väljal [FTUNNUSED]. [arvutada] valitud prognoositav muutuja, 3) prognoosivas komplektis olevate tunnuste ja eelklassifikaatori andmekihid rst-vormingus rasterfailidena. Lisaks saab kasutada interpoleerimiseraldisi (peatükk 7.1). Constud andmebaas prognoosiva komplektiga saadakse Constud õppe tulemusel. Järgnevates ülesannetes kasutatakse eelnevates ülesannetes õpitud muutujaid ja Constud andmebaas peaks seega kaartide loomiseks valmis olema. Prognoositava muutuja valimine väljal [FTUNNUSED].[arvutada] peaks samuti tuttav olema.

Kaardi arvutamiseks peavad prognoosivas süsteemis olevate tunnuste väärtused olema teada kogu uuritava pinna iga üksiku koha kohta. Constud suudab tabelis ATUNNUSED olevate parameetrite järgi arvutada tunnuse andmekihis olevatest piksliväärtustest tunnuse väärtuse igas prognoositavas kohas (väljundrastri pikslis). Tunnuste andmekihid võivad seejuures olla FTP serveris. Andmekihtide leidmisel rakendub Constud põhiaknas sisestatud eesliide ja allika valik: lokaalne või FTP. Arvutatava kaardi suvandite aknas osutatud failidel peavad aga kirjas olema täisnimed ja need failid peavad olema lokaalvõrgu kaudu kättesaadavad.

7.1. INTERPOLEERIMINE

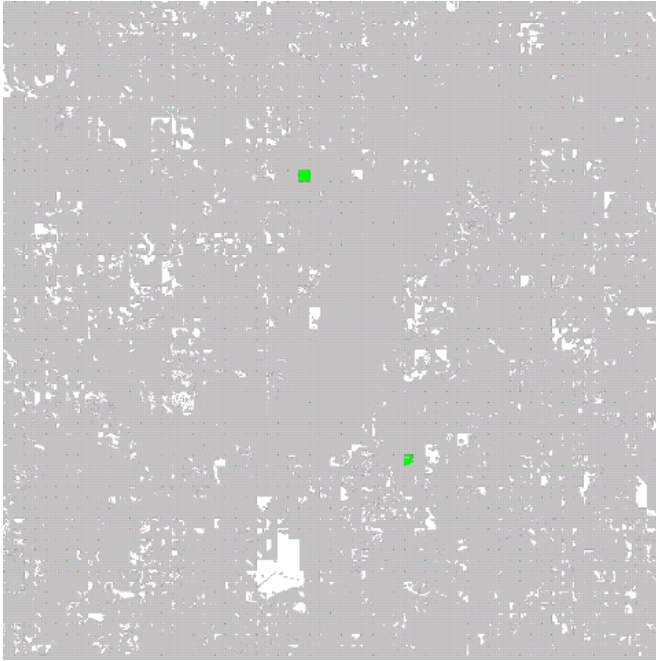
Pikslite suure arvu korral on kaartide moodustamine aeglane, kuna sarnasusele tugineva prognoosimise protseduuri peaks läbi tegema väljundfaili iga piksli kohas. Enamasti on prognoositaval pinnal (maastikul) kompaktsed ühetaolisi osi, kus ei või arvutusi teha suurema vahega kohtades. Maastiku detailsema mosaiiksusega piirkondades ja selgete üleminekute juures oleks hinnangukaarti tarvis lühema sammuga arvutada.

Constud pakub võimalust reguleerida arvutuse sammu ehk ruumilist tihedust. Kasutaja saab määrata algse, kõige pikema sammu pikkuse pikslites. Samm on ridade ja veergude vahe. Kui samm on näiteks kaheksa, siis arvutab Constud prognoosi esmalt iga kaheksanda veeru ja iga kaheksanda rea ristumiskohas. Seejärel kontrollib Constud, kas nurkmiste pikslite kohas prognoositud väärtus on sama (numbrilise muutuja korral kontrollitakse, kas prognoositud väärtuste erinevus on alla lubatud piiri), kas nurkmised pikslid on piisavalt ühetaolised ja interpoleerimiseraldiste kasutamisel, kas nurkmiste pikslite väärtus interpoleerimiskihis on sama.

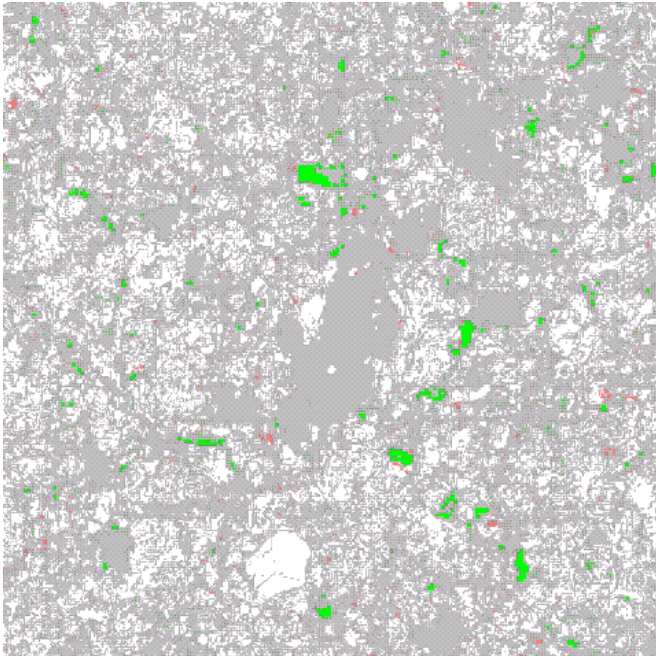
Kui need tingimused on täidetud, siis ei hakka Constud vahepealsete pikslite kohas prognoosi ükshaaval arvutama, vaid omistab nominaalse muutuja korral kogu vahepealsele alale nurkades prognoositud väärtuse. Numbrilise muutuja korral prognoositud väärtus interpoleeritakse. Piirkondades, kus interpoleerida ei õnnestunud ja prognoositud väärtust veel ei ole, kordab Constud hinnangute arvutamist kaks korda väiksema sammuga, kuni kõik pikslid on saanud mingi

tulemuse. Interpoleerimisalgoritmi on kirjeldatud ka tehnilise juhendi peatükis 5.6.3.1 (*Remm ja Linder, 2008*).

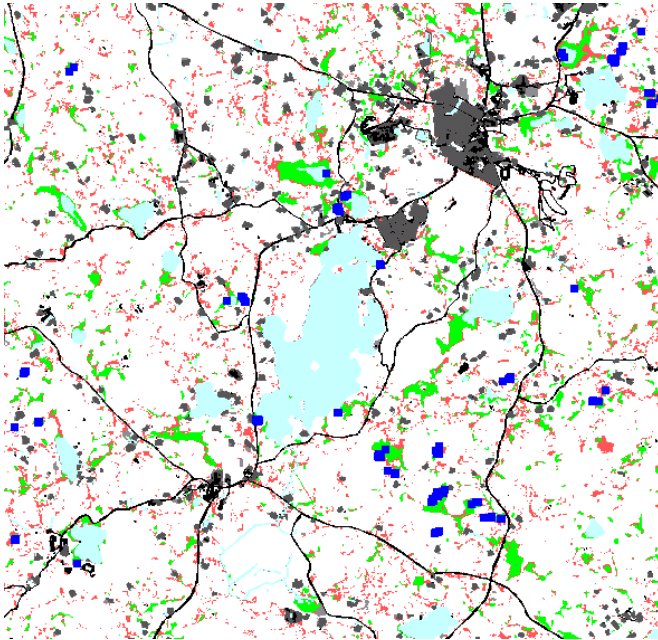
A)



B)



C)



Joonis 7-1. Soo-neiuvaiba hinnangulise esinemiskaardi moodustumine kaardilehel 5434 (Otepää). **A** — interpoleerimise samm 16 pikslit, **B** — samm 4 pikslit, **C** — valmis kaart koos täiendavate teemakihtidega. Hall — määratlemata ala, valge — sarnasem teadaolevatele puudumiskohtadele, roheline — sarnasem teadaolevatele esinemiskohtadele, helepunane — sarnaneb võrdselt nii liigi esinemise kui puudumise kohtadega, helesinine — veekogud, tumesinine — leiukohad, tumehall — hoonestusala, must — suuremad teed. Veekogude ja hoonestusala kohas tõkestas hinnangut **eelklassifikaatorina** kasutatud põhikaardi põhialade kiht. *Remm, et al. (2010).*

Interpoleerimiseraldiste fail peab olema samades piirides ja samasuurte pikslitega, nagu tehtavad kaardidki. Järgnevates ülesannetes läheb interpoleerimiseraldistena tarviskaardilehe 5454 (Elva) rasteriseeritud põhialasid. Kuna interpoleerimiseraldised ei saa olla serveris, siis tuleks eraldiste fail kas Constud veebilehelt muude õpetusandmete juurest või ÖMI digiarhiivist alla laadida.

135. Tõmba oma arvutisse failid 5454_40m.rst ja 5454_40m.rdc Constud veebilehelt või ÖMI digiarhiivi teemakaartide valdkonnast (säilikud FID = 120862 ja 120864). Jäta meelde, millisesse kataloogi failid salvestasid.

7.2. ÜSIKMUUTUJA KAART

Constud põhiakna teadmiste rakenduse osas on valida kahe variandi vahel: üksikmuutuja või komplekshinnang (joonis 7-2). Üksikmuutuja kaardistamisel

võib olla valitud mitu muutujat — hinnangukaardid genereeritakse eraldi kõigile valitud muutujatele. Komplekshinnang eeldab, et väljal [FTUNNUSED].[arvutada] on arvutamiseks valitud vaid ühe FID numbriga muutuja eraldi õpitud kategooriad (funktsioontunnuse tüüp 4).

Hinnangukaardile kantakse prognoositava muutuja väärtused. Constud arvutab koos hinnangukaardiga alati ka kaardi, millel on sarnasused prognoositud väärtuse arvutamisel kasutatud näidistega.

7.2.1. KAHEVÄÄRTUSELINE MUUTUJA

Eelnevates ülesannetes kasutasime kahe erineva kaheväärtuselise muutuja andmestikke. Neist enterobiaasi andmed ei ole asukohakoordinaatidega seotud ja nende alusel Constud süsteemis kaarti teha ei saa. Küll aga saab moodustada soo-neiuvaiba esinemise puudumise sarnasusele tuginevat kaarti.

- ⊙ 136. Lülita väljal [FTUNNUSED].[arvutada] sisse soo-neiuvaiba rida (FID = 2) ja kõik teised read välja.

- ↪ 137. Käivita Constud, vali andmebaas ja Constud põhiakna teadmiste rakenduse osast üksikmuutuja hinnangukaardi arvutus (joonis 7-2). Vajuta nuppu Jätka.

Rasterfaili	<input checked="" type="checkbox"/> Üksikmuutuja hinnang
	<input type="checkbox"/> Komplekshinnang

Joonis 7-2. Üksikmuutuja hinnangukaardi arvutuse valik Constud põhiaknas.

Avaneb kaartide genereerimise aken (joonis 7-3). Akna ülaserava juurest võib lugeda prognoositava tunnuse nime. Selle lugemiseks Constud andmebaasist. Kasutaja sisestada jäävad järgmised suvandid.

Kaardileht — kaardilehe number. Sama numbriga rst-faili otsib Constud iga vajaliku andmekihi kataloogist. Andmekihi kataloogi nimi on kirjas tabelis AKIHID. Andmekihtide puhul rakendub kataloogi nime eesliide (peatükk 1-2).

Piksli külj — piksli külje pikkus samades ühikutes, milles kaardilehtede servade ja vaatluskohtade asukohakoordinaadid. Piksli külj peab olema prognoosifailides ja eraldiste failis sama suur.

Sarnasuse erinevus % — suurim lubatud sarnasuse erinevus interpoleerimisel. Kui erinevus juba arvutatud väärtustega pikslite vahel on suurem, siis interpoleerimist ei rakendata.

Interpoleerimissammu algpikkus — ühikud on pikslid. Interpoleerimissammu vähendatakse kahekordselt igas järgmises korduses.

Hinnangukaardi fail — fail, kuhu salvestatakse prognoositud väärtused.

Näidistega sarnasuse kaardi fail — nimi failile, kuhu salvestatakse iga piksli keskmine sarnasus kasutatud näidistega.

Interpoleerimiseraldiste kasutus — kui valik on väljalülitatud asendis, siis järgmist sisestuskasti ei kuvata.

Interpoleerimiseraldiste fail — nimi failile, kus on interpoleerimiseraldised (peatükk 7.1 ja järgmise suvandi selgitus). Eraldiste fail peab olema samades piirides nagu andmekihid ning sama ridade ja veergude arvuga nagu tulemuste failid.

Kaarti ei arvutata seal, kus piksliväärtus = — tekstiaknasse tuleb sisestada eraldistekihi järgi eiratav väärtus. Joonisel 7-3 on sisestatud number 48, mis tähistab põhikaardi põhiala järvi. Kaardi võiks arvutamata jätta muude veekogude (59), põldude (108) või tootmisõuede (254) kohas. Constud kahjuks ei võimalda ühes andmekihis mitut eiramiskoodi, kuid ühest täiendavast väljamaskimise võimalusest on juttu peatükis 7.2.2. Kui eiratavaid klasse on rohkem kui üks, siis tuleb need klassid eraldiste failis eelnevalt liita. Tänu eiramiskoodile saab eraldistekihti või interpoleerimiseraldiste kihti kasutada arvutatavat ala piiritleva maskina. Nendes võib olla vaid kaks varianti: arvutatav ala ja arvutusest välja jääv ala.

Baite piksli kohta — baitide arv piksli kohta interpoleerimiseraldiste failis. Väärtus üks tähistab ühebaidilisi täisarve (0...255) sisaldavat rastrit, väärtus kaks lühikesi täisarve (-32 768...32 767) sisaldavat rastrit.

Vahetulemuste salvestus — kui on sisse lülitatud, siis salvestatakse iga interpoleerimisvooru tulemus omaette faili. Faili nimesse lisatakse interpoleerimissammu pikkus.

Kihipuudumise hoiatused — valik, kas anda hoiatusteade iga kord, kui Constud ei leidnud mõnda vajalikku andmekihti. Andmekihi puudumisel selle andmekihiga seotud tunnuseid ei kasutata. Sarnasusele tuginevad otsused langetatakse nende tunnuste järgi, mida on igal üksikul juhul võimalik kasutada. Andmekihi puudumise tüüpolukord on, kui mõni õpetusandmete alal olemasolev ja tehisõppes kasutatud andmekiht ei kata kogu prognoositavat ala.

Joonis 7-3. Soo-neiuvaiba hinnangulise esinemise/puudumise arvutuse suvandid kaardilehe 5454 puhul.

- ✓ ↪ **138.** *Sea soo-neiuvaiba esinemise/puudumise kaardi arvutuse suvandid nii, nagu joonisel 7-3. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Pikslite külje pikkuseks vali 40 meetrit, sest eraldistefail on sellise pikslisuurusega. Käivita kaardi arvutus vajutades nupule Arvuta.*


Arvutus võtab aega paar minutit. Kaardi valmides avaneb uuesti Constud põhiaken. Kui väljal [FTUNNUSED].[arvutada] oleks sisse lülitatud rohkem kui üks rida, siis avaneks taas kaardi arvutuse suvandite aken ning nupu *Arvuta* vajutamisel suunduks Constud järgmise muutuja hinnangukaarti tegema.

Kaardi arvutamise aeg sõltub lisaks arvuti protsessori kiirusele veel mitmest faktorist: arvutatavate pikslite arv, andmekihtide hulk, tunnuste hulk ja iga tunnuse arvutuslik keerukus ning andmeühenduse kiirus, kui andmekihid on FTP serveris.

- 🔑 **139.** *Idrisi rst-failide vaatamiseks ja soovi korral konverteerimiseks tuleb igale rst-failile lisada rdc-fail, mille nime põhiosa on sama, mis rst-failil.*

Siin on vähemalt kaks võimalust. Esiteks, kuna eraldistefail ja tulemuste failid on samalaadsed — ridade arv on nendes failides sama, veergude arv on sama ja iga pikslite kohta on salvestatud üks bait — siis piisab faili 5454_40m.rdc (vt ülesanne 135) kopeerimisest ja ümbernimetamisest. Teiseks, kaardilehe 5454 ühebaadises vormingus 40 meetrise sammuga rastrite juurde kuuluvaid rdc-faile saab luua ka alltoodud teksti lihtsasse Notepad vormingus faili kopeerides. Fail tuleb seejärel salvestada ja kirjutada faili nime põhiosaks sama nimi, mis rst-failil ning laiendiks rdc.

file format.: IDRISI Raster A.1
file title.:
data type...: byte
file type...: binary
columns.....: 250
rows.....: 250
ref. system.: plane
ref. units.: m
unit dist...: 1.0000000
min. X.....: 640000.0000000
max. X.....: 650000.0000000
min. Y.....: 6450000.0000000
max. Y.....: 6460000.0000000
pos'n error.: unknown
resolution..: 40.0000000
min. value..: 0
max. value..: 255
display min.: 0
display max.: 255

140. Vaata logifailist, milliste tunnuste poolest sarnasust prognoosikaartide arvutamisel rakendati; ava hinnangulise esinemise/puudumise kaart ja sarnasuse kaart ning ürita mõtestada, et miks saadud kaardid on just sellised, nagu nad on. 

Prognoosikaartide moodustumise vahetulemusi näed failides, mille nime lõpus on interpoleerimissammu number.

Pane tähele, et kaheväärtuselise muutuja prognoosil saadud sarnasus näidistega on otsusekindlus, mitte sarnasus esinemiskohtadega. Sarnasuskaart kirjeldab esinemiskohtadeks prognoositud kohtades sarnasust esinemiskohtade näidistega, puudumiskohtadeks prognoositud kohtades sarnasust puudumiskohtade näidistega.

7.2.2. MULTINOMIAALNE MUUTUJA

Multinomiaalse muutuja näidiseks on taimkatte üksused, mida Constud õppis eristama kasutades põhikaardi põhialasid eelklassifikaatorina (peatükk 4.5). Constud oskab eelklassi kasutust andmebaasist välja lugeda ning kui muutuja puhul on eelklassi kasutus tabelis *FTUNNUSED* sisse lülitatud, kuvab kaardi *suvandite* aknasse eelklassi faili nime sisestamise koha. Lisaks igas eelklassis eraldi näidiste komplekti kasutamisele võimaldab eelklassifikaator osa kaardilehest eirata, nii nagu interpoleerimiseraldisedki. Seega on kahe erineva andmekihi alusel võimalik kaardistatavast kaardilehest mingi osa välja maskida.

- 141. Lülita väljal [FTUNNUSED].[arvutada] sisse rida, kus FID = 3 ja kõik teised read välja.
- 142. Käivita Constud ja vali põhiakna teadmiste rakenduse osast üksikmuutuja hinnangukaardi arvutus (joonis 7-2). Vajuta nuppu Edasi.
- 143. Sea taimkattekaardi arvutuse suvandid nii, nagu joonisel 7-4. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Piksli külje pikkuseks vali 40 meetrit, sest eraldistefail on sellise pikslisuurusega. Käivita kaardi arvutus vajutades nupule Arvuta.

Interpoleerimiseraldised ja eelklassifikaator on seekord samad, mis ei pea olema reegel. Interpoleerimiseraldiste järgi maskitakse seekord välja järved, eelklassifikaatori järgi põllud.

- 144. Genereeritud failide vaatamiseks ja konverteerimiseks tuleb lisada rdc-fail, mille nime põhiosa on rst-failiga ühine; rdc-faili sisu peaks olema sama, mis ülesanne 139 juures.
- 145. Vaata logifailist, milliste tunnuste poolst sarnasust prognoosikaartide arvutamisel rakendati, ava hinnanguline taimkattekaart ning sarnasuse kaart ning ürita mõtestada, et miks saadud kaardid on just sellised, nagu nad on.

Prognoosikaartide moodustumise vahetulemusi näed failides, mille nime lõpus on interpoleerimissammu number.

Joonis 7-4. Hinnangulise taimkattekaardi suvandid kaardilehe 5454 puhul.

7.2.3. NUMBRILINE MUUTUJA

Numbrilise muutuja puhul lisandub interpoleerimise tingimuste juurde suurim lubatud prognoositud väärtuste erinevus kaardistatava muutuja ühikutes. **Nominaalse muutuja** puhul oli interpoleerimine üle ruudukujulise vahemiku lubatud, kui prognoositud üksus on igas nurgas sama.

Sademetes õppes (ülesanne 62) olid tunnused seotud universaalsete kujuteldavate andmekihtidega. Constud õppeks ja tabelisse prognoosimiseks sellest piisas, prognoosikaardi jaoks aga mitte, sest kaardi moodustamiseks on tarvis **argumenttunnuste** väärtusi igas kaardistatava ala punktis. Tinglikest kujuteldavatest andmekihtidest tunnuseid arvutada ei saa. Seetõttu tuleks andmekihtide tabelisse lisada sademetes õppes kasutatud tunnustele tegelikult vastavate andmekihtide kirjed ning muuta tunnustega seotud andmekihtide numbrid tabelis **ATUNNUSED**. Senini on kaardilehtede tabelis **PK20t** vaid Eesti põhikaardi lehtede numbrid ja servakoordinaadid, Baltimaade sademetes andmekihtide failid katavad aga kogu Baltikumi. Järgmised ülesanded näitavad, et Constud abil loodud kaardid ei pruugi olla seotud Eesti põhikaardi kaardilehtede nomenklatuuriga ega põhikaardi projektsiooniga.

146. Kopeeri täiendavate andmekihtide 34...40 kirjeldused näidisandmete andmebaasi tabelist Lisakihid tabelisse AKIHID Constud andmebaasis.

Oluline on, et andmekihtide KID numbrid, piksli suurused ja kataloogid oleksid õiged. Kataloogid näitavad iga andmekihi paiknemist ÖMI digiarhiivis. Väljaspool Tartu ülikooli arvutivõrku olevatel kasutajatel tuleb andmekihid ise hankida ja seadistada.

147. Lisa tabelisse PK20t kaardileht numbriga 1 ja servakoordinaatidega $\min-x = 304000$, $\min-y = 5968000$, $\max-x = 767000$, $\max-y = 6620000$.

Koordinaatsüsteem ei ole siin Lambert-Est, vaid TM-Balti, aga Constud kasutab vaid ristkoordinaate ja koordinaatsüsteemiga ei tegele. Kaardilehe number 1 on tinglik, aga peab täpselt vastama andmekihi failidele.


148. Ava tabelid ATUNNUSED ja AKIHID ning paranda tunnuste 51...57 andmekihtide numbrid tabeli ATUNNUSED väljal KID viies vastavusse tunnuse nime ja andmekihi kataloogi.

Ära muuda ühegi andmekihi kataloogi nime. Kui oled tegutsenud selle juhendi järgi, siis peaksid uued andmekihid samas järjekorras olema, mis tunnusedki.



149. Tõmba Constud veebilehelt või ÖMI digiarhiivi CS-kihtide valdkonnast oma arvutisse fail **Baltimaa_koodiga_o.rst** ja fail **Baltimaa_koodiga_o.rdc** (digiarhiivi säilikud FID = 120873 ja 120872). Jäta meelde, millisesse kataloogi faili salvestasid.

See fail sobib hästi Baltimaade maskiks. Selles failis on kolme Balti riigi territooriumi kohas piksliväärtused = 0, mujal = 255.

- 150. Lülita väljal [FTUNNUSED].[arvutada] sisse Balti sademete rida (FID = 4) ja kõik teised read välja.



 151. Käivita Constud ja vali põhiakna teadmiste rakenduse osast üksikmuutuja hinnangukaardi arvutus (joonis 7-2).

  152. Lülita välja valik Ruut-kaardilehed. Vajuta nuppu Jätka.


  153. Sea Balti sademete kaardi arvutuse suvandid nii, nagu joonisel 7-5. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Piksli külje pikkuseks vali 1000 meetrit, sest eelklassifitseeriva maski failis on selle suurusega pikslid. Käivita kaardi arvutus vajutades nupule Arvuta.

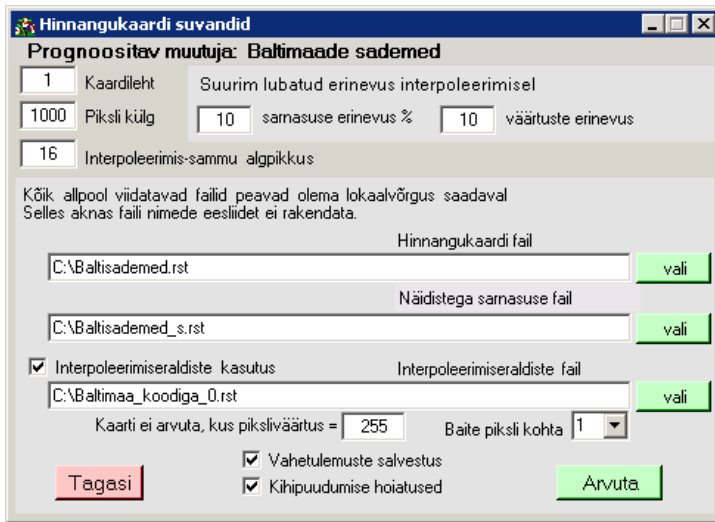
Constud ei anna õiget akent ette ehk ta ei anna väärtuste erinevus akent. Samuti ei näita prognoositavat muutujat.

Eelklassifikaator jääb seekord kasutamata, sest seda ei olnud ka õppes. Interpoleerimiseraldisteks on ülesande 149 juures alla laaditud Baltimaade mask, kus muu ala on piksliväärtusega 255. Vahetulemuste salvestuse võid ka välja lülitada.

-   154. Moodustatud failide vaatamiseks ja konverteerimiseks tuleb igale failile lisada rdc-fail, mille nime põhiosa on sama, mis rst-failil.

Sarnasusekaardi rdc-faili saab faili *Baltimaa_koodiga_0.rdc* kopeerides ja ümber nimetades. Prognoositud sademete faili vaatamiseks tuleb rdc-faili sisu veidi muuta, sest numbrilise muutuja prognoosifail on neljabaidise murdarvu vormingus. Asenda prognoositud sademete rdc-failis tekst: „data type : byte“ => „data type : real“. Sarnasuskaardi rdc-failis jääb andmetüübiks byte.

-  155. Vaata logifailist, milliste tunnuste poolest sarnasust prognoosikaartide arvutamisel rakendati, ava hinnanguline sademete hulga kaart ja näidistega sarnasuse kaart ning ürita mõtestada, et miks saadud kaardid on just sellised, nagu nad on.



Hinnangukaardi suvandid

Prognoositav muutuja: Baltimaade sademed

1 Kaardileht Suurim lubatud erinevus interpoleerimisel

1000 Piksli külg 10 samasuse erinevus % 10 väärtuste erinevus

16 Interpoleerimis-sammu algpikkus

Kõik allpool viidatavad failid peavad olema lokaalvõrgus saadaval
Selles aknas faili nimede eesliidet ei rakendata.

Hinnangukaardi fail
C:\Baltisademed.rst

Näidistega samasuse fail
C:\Baltisademed_s.rst

Interpoleerimiseraldiste kasutus Interpoleerimiseraldiste fail
C:\Baltimaa_koodiga_0.rst


Kaarti ei arvuta, kus piksliväärtus = 255 Baite piksli kohta 1

Vahetulemuste salvestus Kihpuudumise hoiatused



Joonis 7-5. Hinnangulise sademetekaardi moodustamise suvandid.

7.2.4. NOMINAALSE MUUTUJA ÜKSUS

Kui **nominaalse muutuja** eraldi õpitud üksuse kaardistamine käivitatakse Constud põhiakna üksimuutuja hinnangu valiku (joonis 7-2) kaudu, siis käsitleb Constud seda kaheväärtuselise muutuja prognoosikaardi moodustamise ülesandena. Seega iga eraldi õpitud üksuse esinemise/puudumise kohta arvutatakse eraldi kaart.

156. Lülita väljal [FTUNNUSED].[arvutada] sisse taimkatte eraldi õpitud üksused aruniit, sööt ja põld ning kõik teised read välja. Märgi eraldi arvutatud üksuste ([FTUNNUSED].[FID] = 5) eelklassifikaatoriks taas 13. Eelklassifikaatori välja tühjendasid ajutiselt täites ülesannet 129. 

157. Käivita Constud ja vali põhiakna teadmiste rakenduse osast üksikmuutuja hinnang (joonis 7-2). Vaata, et valik Ruut-kaardilehed oleks sisse lülitatud. Vajuta nuppu Jätka.  

158. Sea kaardi arvutuse suvandid nii, nagu on joonisel 7-6. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Piksli külje pikkuseks vali 40 meetrit, sest eelklassifitseeriva maski failis on selle suurusega pikslid. Käivita kaardi arvutus vajutades nupule Arvuta.  

Pane tähele, et prognoositava muutuja nimeks on esimene valitud üksus, praegusel juhul aruniit. Õige on kasutada sama **eelklassifikaatorit**, mida kasutati selle muutuja õppes. Kui õppesse eelklassifikaatorit ei kaasatud, siis ei ole põhjust seda ka prognoosikaardi arvutamisse kaasata. Interpoleerimiseraldiste kasutuse ja vahetulemuste salvestuse võid ka välja lülitada. Joonise 7-6 järgi jäetakse hinnangud arvutamata järvede (kood 48) ja tootmisõuede kohas (kood 254). Võid eiratavateks

7.3. KOMPLEKSHINNANG

Hinnanguliste kaartide arvutamine **nominaalsele muutujale** iga üksuse omaette õppe järel erineb arvutuslikult hinnangulise kaardi teistest variantidest ja on seetõttu Constud põhiaknas omaette valik nimega kompleks hinnang (joonis 7-7). Komplekshinnangu valikut saab kasutada vaid nominaalse muutuja eraldi õpitud üksuste puhul (**funktsioontunnuse** tüüp 4). Komplekshinnangu valimise ja nupule *Jätka* vajutamise järel avaneb hinnangulise kaardi moodustamise aken, millel on kolm varianti (joonis 7-8).

Esimese variandi puhul leitakse nominaalse muutuja väljal [FTUNNUSED].[arvutada] valitud üksuste hulgast iga koha tunnustega kõige sarnasem üksus ning keskmine sarnasus prognoosi arvutamisel kasutatud näidistega (joonis 7-8A). Teise variandi puhul arvutatakse iga koha sarnasus selles kohas ette antud üksusega (kaardistatud üksusega) (joonis 7-8B). Kolmanda variandi puhul arvutatakse nii sarnasus etteantud üksusega kui ka prognoositud üksus (joonis 7-8C).

Kuna eraldi õpitud üksuste puhul on igal üksusel oma tunnuste komplekt, siis ei saa prognoositud üksust ühekorruga arvutada. Constud arvutab igas kohas sarnasuse iga kaasatud üksusega. Prognoositud üksuseks saab see, millega sarnasus on suurim.

Komplekshinnangut saab kasutada vaid nominaalse
muutuja puhul, mille üksusi on Constud eraldi õppinud

Rasterfaili

- Üksikmuutuja hinnang
- Komplekshinnang

Joonis 7-7. Eraldi õpitud üksustega nominaalse muutuja hinnangukaardi arvutuse valik Constud põhiaknas.

A)

Hinnangukaardi suvandid

Proгноositav muutuja: FID= 5

5454 Kaardileht Suurim lubatud erinevus interpoleerimisel

40 Piksli külg 10 sarnasuse erinevus %

8 Interpoleerimis-sammu algpikkus

Kõik allpool viidatavad failid peavad olema lokaalvõrgus saadaval
Selles aknas faili nimede eesliidet ei rakendata.

Hinnangukaardi fail
C:\5454taimkatte eraldi õpitud üksused.rst vali

Näidistega sarnasuse fail
C:\5454sarnasus prognoositud üksusega.rst vali

Interpoleerimiseraldiste kasutus Interpoleerimiseraldiste fail
C:\5454_40m.rst vali
Kaarti ei arvuta, kus piksliväärtus = 48 Baite piksli kohta 1

Prognoositava muutuja eelklassi kasutus Prognoositava muutuja eelklasside fail
C:\5454_40m.rst vali
Kaarti ei arvuta, kus piksliväärtus = 254

Sarnasus kaardistatuga

Vahetulemuste salvestus

Tagasi Arvuta

B)

Hinnangukaardi suvandid

Proгноositav muutuja: FID= 5

5454 Kaardileht Suurim lubatud erinevus interpoleerimisel

40 Piksli külg 10 sarnasuse erinevus %

8 Interpoleerimis-sammu algpikkus

Kõik allpool viidatavad failid peavad olema lokaalvõrgus saadaval
Selles aknas faili nimede eesliidet ei rakendata.

Etteantud koodiga sarnasuse fail
C:\5454sarnasus etteantud taimekatteüksusega.rst vali

Interpoleerimiseraldiste kasutus Interpoleerimiseraldiste fail
C:\5454_40m.rst vali
Kaarti ei arvuta, kus piksliväärtus = 48 Baite piksli kohta 1

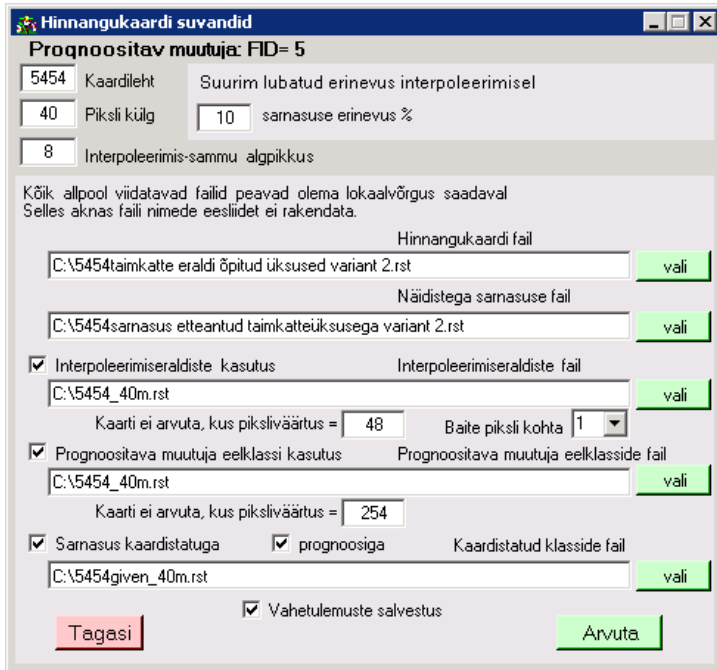
Prognoositava muutuja eelklassi kasutus Prognoositava muutuja eelklasside fail
C:\5454_40m.rst vali
Kaarti ei arvuta, kus piksliväärtus = 254

Sarnasus kaardistatuga prognoosiga Kaardistatud klasside fail
C:\5454given_40m.rst vali


Vahetulemuste salvestus



Tagasi Arvuta



C)



Joonis 7-8. Eraldi õpitud üksustega **nominaalse muutuja** hinnangukaardi **suvandid**. **A** — prognoositud üksuse ja **näidistega sarnasuse** kaart, **B** — sarnasus igas kohas etteantud üksusega, **C** — sarnasus igas kohas etteantud üksusega ja prognoositud üksus.

162. Lülita väljal [FTUNNUSED].[arvutada] sisse kõik taimkatte eraldi õpitud üksused (FID = 5) ja kõik teised read välja. 

163. Käivita Constud ja vali põhiakna hinnangute arvutuse osast komplekshinnang   (joonis 7-7). Vaata, et valik Ruut-kaardilehed oleks sisse lülitatud. Vali andmebaas ja vajuta nuppu Jätka.

164. Sea kaardi arvutuse suvandid nii, nagu on joonisel 7-8A. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis.   Piksli külje pikkuseks vali 40 meetrit, sest eelklassifitseeriva maski failis on selle suurusega pikslid. Käivita kaardi arvutus vajutades nupule Arvuta.


Pane tähele, et prognoositava muutuja nime kaardi suvandite aknas ei näidata, sest seda ei ole andmebaasis kirjas. Andmebaasis on eraldi õpitud üksustega muutuja FID number ja iga üksuse nimi, muutuja kui terviku nime tabelis FTUNNUSED ei ole.

Interpoleerimiseraldiste kasutuse ja vahetulemuste salvestuse võid ka välja lülitada. Õige on kasutada sama **eelklassifikaatorit**, mida kasutati selle muutuja õppes. Kui


õppesse eelklassifikaatorit ei kaasatud, siis ei ole põhjust seda ka prognoosikaardi arvutamisse kaasata. Joonise 7-8A järgi jäetakse hinnangud arvutamata järvede (kood 48) ja tootmisõuede kohas (kood 254). Võid eiratavateks põhialadeks valida ka mõned muud eraldistefailis *5454_40m.rst* olevad põhikaardi põhialad. Põhialade nimed ja koodid on näidisandmete andmebaasi *Constud_6ppe_andmed.mdb* tabelis *Põhialad*.

Kaartide valmides avaneb taas Constud põhiaken, millest saab nupu *Jätka* vajutamisega uuesti avada hinnangukaartide arvutamise suvandite akna.


Järgmisena moodustame kaardi, mis näitab sarnasust etteantud kaardil samas kohas oleva üksusega. Etteantud kaardina saab kasutada kas ülesande 143 juures arvutatud taimkattekaarti või faili *5454_given_40m.rst*, mille võib leida muude näidisandmete hulgast Constud veebilehel ja **ÖMI digiarhiivist** (säilik 120885 ja 120886 teemakaartide osas).

 **165. Hangi Constud veebilehel olevate näidisandmete hulgast või ÖMI digiarhiivist (säilik 120885 ja 120886 teemakaartide osas) fail *5454_given_40m.rst*.**

Selles failis on äärmiselt lihtsustatud teoreetiline taimkatte mudel, kus kõik põhikaardi põllud on klassifitseeritud söödiks, kõik metsad palumännikuks ja kõik rohumaad aruniiduks. Kõik muud alad on piksliväärtusega 0. Arvutatav sarnasuskaart peaks väiksemate piksliväärtustega näitama alasid, mis ei sarnane nende ette antud üksustega; sarnasus sada või saja lähedal näitab, et koht võib tõesti etteantud klassi kuuluda. Järgnevates harjutustes kaardistame iga koha sarnasust sellel kaardil ette antud üksusega. Järgnevad kaks harjutust erinevad vaid prognoosikaardi arvutamise poolest. Esimesel juhul arvutatakse vaid sarnasuskaart, teisel juhul nii sarnasuskaart kui ka prognoosikaart.

 **166. Sea kaardi arvutuse suvandid nii, nagu on joonisel 7-8B. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Piksli külje pikkuseks vali 40 meetrit, sest eelklassifitseeriva maski failis on selle suurusega pikslid. Vahetulemused võid ka salvestamata jätta. Käivita kaardi arvutus vajutades nupule Arvuta.**

Arvutatakse vaid sarnasus kaardil oleva üksusega ehk etteantud üksusega.

 **167. Sea kaardi arvutuse suvandid nii, nagu on joonisel 7-8C. Kataloogid kohanda vastavaks failide paiknemisega oma arvutis. Piksli külje pikkuseks vali 40 meetrit, sest eelklassifitseeriva maski failis on selle suurusega pikslid. Vahetulemused võid ka salvestamata jätta. Käivita kaardi arvutus vajutades nupule Arvuta.**

Arvutatakse nii sarnasus etteantud üksusega kui ka prognoositud üksus. Prognoosi arvutus võtab märksa rohkem aega kui vaid sarnasuskaardi arvutus.

168. Lisa ülesande 164, 166 ja 167 tulemusel arvutatud kaartide lõppvariantidele rdc-failid.

Lõppvariandid on need, mille nime lõpus ei ole interpoleerimissammu numbrit. Rdc-failid saad kopeerides ja ümber nimetades mõnda olemasolevat rdc-faili, mis on kaardilehe 5454 kohta, mis kirjeldab ühebaadist Idrisi rastervormi ja 40 meetrise küljega piksleid.

169. Võrdle ülesannete 164 ja 167 tulemusel moodustatud hinnangulisi taimkattekaarte.

Selle juhendi järgi toimides peaks nende kaardifailide nimed olema: *5454taimkatte eraldi õpitud üksused.rst* ja *5454taimkatte eraldi õpitud üksused variant 2.rst*.

170. Võrdle ülesannete 164, 165 ja 167 tulemusel moodustatud sarnasuskaarte.

7.4. ASENDUSTUNNUSED

Indeksite arvutamisel ja muutujate õppes kasutatakse tunnuseid, mis on sisse lülitatud väljal *[ATUNNUSED].[F...]*. Tunnuste asendamiseks või dubleerimiseks puudub siin reeglina vajadus. Kaartide arvutamisel aga annab olulist ajasäästu keerukamate ja korduvalt kasutatavate tunnuste täiendavateks andmekihtideks valmis arvutamine. Rasterfailidesse valmis arvutatud tunnused moodustavad igaüks omaette andmekihi, mis tuleb registreerida tabelis *AKIHID*. Tunnuste andmekihtidena valmis arvutamiseks saab kasutada indeksite rasterfaili arvutamise funktsiooni (peatükk 3.2).

Valmis arvutatud tunnused võivad esiteks olla tabelis *ATUNNUSED* määratletud asendustunnustena. Sel juhul on sama tunnus *argumenttunnuste* tabelis kirjas kaks korda (joonis 7-9). Asendustunnus peab kindlasti olema suurema AID numbriga kui originaaltunnus. Asendustunnuseid kasutatakse vaid kaartide arvutamisel. Asendustunnus peaks olema võimalikult lihtsasti arvutatav. Parim variant numbrilise väljundiga tunnuste puhul on märkida indeksi koodiks 102 (keskväärtus, selle arvutus on kõige kiirem), valimiks 1 või tühi lahter (pikslite juhuvalikut ei rakendata) ja raadiuseks null või jäta raadiuse lahter tühjaks (loetakse vaid fookuses olev piksliväärtus). Nominaalse väljundiga on vaid ühel või teisel viisil arvutatud *mood*, mille asendamiseks sobib indeks 2 (esimese astme mood).

Teine tunnuste asendamise võimalus on muuta kaartide arvutamise eel tabelis *ATUNNUSED* iga valmis arvutatud tunnuse andmekiht, indeks ja indeksi arvutamise raadius. Tunnuse kirjelduse muutmise puudus on, et kaardistatava muutuja õppe jätkamiseks peaks tabeli *ATUNNUSED* tagasi kohandama või siis muutma

ka tunnuseid õpetusandmetes. Mida rohkem muudatusi, seda suurem risk kuskil eksida.

ATUNNUSED					
AID	nimi	KID	indeksID	radius	asendab
67	nõlva suund	3	109	500	
68	nõlva suund	50	102		67
69	kaalutud põhiala	1	11	1000	
70	kaalutud põhiala	51	1		69

Joonis 7-9. Õppes kasutatud tunnused 67 ja 69 ja nende asendustunnused 68 ja 70 tabelis *ATUNNUSED*. Kaartide moodustamisel ei arvutata üldistatud nõlva suunda pinnakõrguse andmekihist KID = 3 iga väljundpikslit kohas 500 ühiku raadiuses, vaid loetakse see tunnus arvutatavas kohas andmekihis KID = 50 olevalt piksliväärtuselt. Samuti ei arvutata kaugusega kaalutud sagedasimat põhiala kihist 1, vaid loetakse see valmis arvutatud kihist KID = 51.



KÜSIMUSED

1. Kas **Constud** kasutab interpoleerimiseraldiste andmekihti, kui interpoleerimise sammuks on märgitud 1?
2. Mis juhtub, kui **Constud** süsteemis prognoosikaardi kaardilehel mõnda andmekihti ei ole?
3. Kas sama **nominaalse muutuja** erinevate üksuste hinnangulistel levikukaartidel võib sama koht olla märgitud tõenäoliseks esinemiskohaks rohkem kui ühe üksuse puhul?
4. Kas ülesannete 164 ja 167 tulemusel moodustatud hinnangulised taimkattekaardid on erinevad? Kui jah, siis mille poolest? Selle juhendi järgi toimides peaks nende nimed olema: *5454taimkatte eraldi õpitud üksused.rst* ja *5454taimkatte eraldi õpitud üksused variant 2.rst*.
5. Mitut erinevat taimkatteüksust kujutab ülesande 164 järgi moodustatud kaart *5454taimkatte eraldi õpitud üksused.rst*? Mis piirab üksuste arvu sellel kaardil?
6. Kas ülesannete 164 ja 167 tulemusel moodustatud sarnasuskaardid failides *5454sarnasus etteantud taimkatteüksusega.rst* ja *5454sarnasus etteantud taimkatteüksusega variant 2.rst* on erinevad?
7. Kas sarnasuskaart etteantud üksustega ja sarnasuskaart prognoosil kasutatud **näidistega** on erinevad? Kui jah, siis kummas on sarnasused reeglina suuremad ja miks?
8. Kas **Constud** abil arvutatav hinnanguline kaart võib olla ebakorrapärase kujuga?
9. Miks tuli **Constud** abil Baltimaade sademete kaardi arvutamisel kaardilehe numbriks märkida 1?

LÕPPSÕNA

Constud pakub nii palju erinevaid võimalusi, et kõikide kasutamisanvariantide kohta ei ole mõtet eraldi ülesannet koostada. Mõni omadus ja puudus võib jääda ka kasutaja enda avastada. Näiteks ei ole selles õpikus näidisülesannet sama nominaalse muutuja üksuste omavahelise sarnasuse seadistamise kohta ega numbriliste tunnuste tabelisse valmis arvatatud standardhälvete kasutamisest ja selle mõjust tulemustele.

Tarkvara arendatakse seni, kuni on kasutajaid. Kui areneb tarkvara, siis peab muutuma ka selle kirjeldus ja kasutamise õpik. Õpikut aitab täiendada iga asjalik soovitus ja kriitiline märkus. Märkusi ja ettepanekuid saab saata veebilehelt <http://digiarhiiv.ut.ee/Constud2/Kasutamisopetus.aspx>.

Selle õpiku koostamise käigus kontrolliti, parandati ja muudeti **Constud** tarkvara päris suurel määral. Keeruka süsteemi muutmisega kaasneb paraku risk, et uuendus algkoodi ühes osas toob kaasa tõrke mingis teises kohas või mingite teiste varem kontrollitud suvandite kasutamisel. Autorid loodavad kasutajate mõistvale suhtumisele ja kaasabile **Constud** süsteemi arendamisel. Abi on igast teatest, et süsteem ei toiminud oodatud viisil, samuti täiendavate soovide teada andmisest.

VIIDATUD KIRJANDUS

Allouche, O., Tsoar, A., Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223–1232.

Congalton, R.G., Green, K. 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Lewis Publishers.

Remm, K. 2004. Case-based predictions for species and habitat mapping. *Ecological Modelling*, 177(3-4), 259–281.

Remm, K., Linder, M. 2008. *Tehisõppesüsteemi Pidevstudium kirjeldus. Versioon 28.11.08*. Käsikiri Tartu Ülikooli Ökoloogia ja Maateaduste Instituudis ja veebilehel <http://digiarhiiv.ut.ee/Constud2/Doc/Pidevstudium.pdf>.

Remm, K., Linder, M., Proosa, H. 2010. *Description of the Software System Constud 2. Version 27.08.10*. Chair of Geoinformatics and Cartography, University of Tartu (http://digiarhiiv.ut.ee/Constud2/Doc/Software_system_Constud.pdf).

Remm, K., Remm, L. 2009. Similarity-based large-scale distribution mapping of orchids. *Biodiversity and Conservation*, 18(6), 1629–1647.

Remm, K., Remm, M. 2010. Geographical aspects of enterobiasis in Estonia. *Health & Place*, 16, 291–300.

Remm, M., Remm, K. 2008. Case-based estimation of the risk of enterobiasis. *Artificial Intelligence in Medicine*, 43(3), 167–177.

Tamm, T., Remm, K. 2009. Estimating the parameters of forest inventory using machine learning and the reduction of remote sensing features. *International Journal of Applied Earth Observation and Geoinformation*, 11(4), 290–297.

LISA
1280

LISA 1. MÕISTED

Andmebaasiobjekt — andmebaasi loogilise ülesehituse osis; selles õpikus relatsioonilise andmebaasi tabel, päring või vaade.

Andmebaasivaade — salvestatud päring või muu reeglistik, mis määrab andmete valiku ja kasutajale esitamise vormi.

Andmeväli — koht andmete hoidmiseks; veerg andmete tabelkujul esitamisel.

Andmekaevandamine — reeglipärasuse otsimine suurtest andmehulkadest kombineerides statistika, tehiseõppe ja andmebaasihalduse meetodeid.

Argumenttunnus ehk **seletav tunnus**, tunnus mille väärtustest sõltuvad funktsioontunnuse väärtused.

Atribuut — objektiga lahutamatu seotud omadus või täiend.

Determineeritud protsess — toiming, mis sama algseisundi puhul jõuab alati sama lõpptulemuseni, kuna ei sisalda juhuslikku komponenti. Indeksite arvutus (kui ei kasutata juhuvalimeid) ja teadmiste rakendus süsteemis **Constud** on determineeritud, **Constud tehiseõpe** ei ole.

Diskreetne muutuja — muutuja, mille võimalike väärtuste hulk on loendatav. Võib olla sisuliselt järjestatav või mitte.

Divisor — ehk jagaja, **Constud** süsteemis suvaline arv, millega jagatakse tunnuse esialgsed väärtused.

Eelklassifikaator — tunnus, mille järgi saab andmestiku allosadeks jagada ja iga osa eraldi käsitleda.

Eraldis — uuritava ruumi mingi tunnuse või reegli alusel piiritletud sidus osa. Tavaliselt polügoon kaardil.

Funktsioontunnus (sõltuv tunnus) — tunnus, mille väärtused sõltuvad argumenttunnuste ehk seletavate tunnuste väärtustest.

Iteratsioon — kordus. Iteratiivne protsess on kordustega toiming.

Kernel — tuumik, **Constud** indeksite arvutamise osas liikuv aken, mille sees olevate pikslite väärtusi kasutatakse **ruumimustri** indeksite arvutamisel.

Positiivse ruumilise autokorrelatsiooni puhul kalduvad lähestikku paiknevad kohad sama tunnuse poolest sarnasemad olema kui kaugemad kohad.

Ruumimuster — objektide või tunnuse väärtuste ruumis paiknemise viis.

Ruutkeskmise viga — numbrilise muutuja prognoositud või hinnatud väärtuste ning vaadeldud väärtuste vahelise erinevuse näitaja, arvutuslikult ruutjuur keskmisest ruuthälbest.

Seletav tunnus (argumenttunnus) — tunnus, mille väärtustest funktsioontunnuse väärtused sõltuvad.

SRTM (*Shuttle Radar Topographic Mission*) — rahvusvaheline projekt globaalse kõrgusmudeli loomiseks.

Suvand — tarkvara seadistus, mida kasutaja saab oma suva järgi muuta.

Sõltuv muutuja (funktsioontunnus ehk prognoositav tunnus) — muutuja, mille väärtused on seatud sõltuvusse seletavate tunnuste ehk argumenttunnuste väärtustest.

Standardhälve — ruutjuur muutuja hajuvusest.

Tehisõpe — protsess, millega iseõppiv süsteem täiendab oma teadmisi. Tarkvaras Constud aina korduv toiming võimalikult õigeid hinnanguid andvate kaalude leidmiseks tunnustele ja näidistele. Eelmiste korduste tulemused mõjutavad järgmiste korduste algseisu.

Tõesuskriteerium — kindla eeskirja järgi arvatav koondnäitaja, mille alusel võrreldakse erineval viisil või erinevates valimites saadud hinnangute ja vaatlustulemuste vastavust.

Vaatlus — üldmõiste andmetöötluses. Üks rida objekt-tunnus maatriksis, mille veergudes on tunnused ja ridades on tunnustega kirjeldatud objektid (isikud, mõõtmistulemused, katsed, juhtumid vms).

Vabadusastmete arv — mudeli sõltumatute osade arv. Üldjuhul omavahel sõltumatute vaatluste arv miinus mudelis kasutatud tunnuste arv.

Valimi maht — vaatluste arv valimis.

Vastavus — kaudsete hinnangute ja otseste vaatluste kokkulangevus.

Vektorkuju — ruumiandmete organiseerimise vorm, kus asend kirjeldatakse punktide, joonte ja pindadena ning nende kuju ja vormi määravate matemaatiliste funktsioonidega, mitte väärtuste regulaarse võrgustikuna.

Õmblusteta andmekiht — andmete hoidmise vorm, milles sisemise ruumilise jaotuse piirid (kaardilehtede piirid) ei ole nähtavad.

Õpetusandmed — andmed, mille abil sobitatakse (kalibreeritakse) mudeli või muu prognoosiva süsteemi parameetrid (arvulised omadused, mida saab muuta).

Õppevastavus — prognooside ja vaatluste vastavus õppes või mudeli kalibreerimisel kasutatud valimis.

Õppekordus (õppevoor) — tarkvaras **Constud** tunnustele ja **näidistele** kaalude sobitamise üks kordus, mis lõpeb saadud kogemuse salvestamisega **Constud** andmebaasi. Õppevoor sisaldab kaalude sobitamise erinevaid mooduseid; iga mooduse sees kulgeb samuti kordustega toiming.

ÖMI digiarhiiv — Tartu Ülikooli Ökoloogia ja Maateaduste Instituudi digitaalsete kaardi- ja kaugseireandmete arhiiv.

LISA 2. CONSTUD EESTIKEELSE VERSIOONI ANDMETABELID

Lisa 2.1. Vajalikud väljad tabelis *FTUNNUSED*

Välja nimi	Vorming	Kommentaar
arutada	jah/ei	Funktsioontunnuse valik
FID	bait	funktsioontunnuse number
antudkood	bait	Klassikood multinominaalse tunnuse klasside ükshaaval käsitlemisel
LogID	4-baidine täisarv	Seni parima tunnuste ja näidiste komplekti tunnusnumber logitabelis
eelklass	bait	Eelklassifikaatorina kasutatava tunnuse number. Tühi väli, väärtused 0 ja 255 tähistavad eelklassifikaatori
nimi	tekst	Funktsioontunnuse nimi
ftüüp	bait	Funktsioontunnuse tüüp
osatunnuseid	bait	Mitmemõõtmelise funktsioontunnuse osatunnuste arv
sumsimmax	4-baidine murdarv	Otsitava kogusarnasuse algväärtus (vajalik vaid tehisõppes)
andmepäring	tekst	Andmebaasiobjekt, millest loetakse õpetusandmed ja näidised ning näidiste kaalud ja aktuaalsused
vmaht	4-baidine täisarv	Õpetusvaatluste ligikaudne arv vaatluste juhuvalimis (vajalik vaid tehisõppes). Väärtused >2000 asenduvad valimi mahuga 2000.
tmaht	4-baidine täisarv	Tunnuste ligikaudne arv tunnuste juhuvalimis (vajalik vaid tehisõppes)
kvmaht	4-baidine täisarv	Kontrollvalimi maht (vajalik vaid tehisõppes). Väärtused >5000 asenduvad valimi mahuga 5000.
nullkaugus	4-baidine murdarv	Sarnasuse nullkaugus (vajalik vaid tehisõppes)

Otsitav kogusarnasus on prognoositava vaatluse ja näidiste vahelise summaarse sarnasuse väärtus, mille järel lõpetatakse sarnasuselt järgmiste näidiste otsimine.

Tehisõppe osas vähendatakse vaatlustevahelist sarnasust proportsionaalselt vaatlustevahelise vahemaa pöördväärtusega. Kui vaatluste vahemaa ei ületa nullkaugust, siis nendevaheline sarnasus nullitakse, kui vaatluste vahemaa võrdub kahekordse nullkaugusega, siis vähendatakse sarnasust 50% võrra.

Lisa 2.2. Vajalikud väljad tabelis ATUNNUSED

Välja nimi	Vorming	Kommentaar
AID	2-baidine täisarv	Argumenttunnuse tunnusnumber
nimi	tekst	Tunnuse nimi
KID	bait	Andmekihi tunnusnumber tabelis AKIHID
indeksID	bait	Indeksi tunnusnumber tabelis INDEKSID
radius	4-baidine murdarv	Tunnuse arvutamise akna välisraadius (vajalik vaid indeksite ja kaartide arvutamisel)
sradius	4-baidine murdarv	Tunnuse arvutamise akna siseradius (vajalik vaid indeksite ja kaartide arvutamisel)
valim	4-baidine murdarv	Pikslite juhuvalimi maht vahemikus 0...1 tunnuse arvutamisel (tehisõppes ei ole vajalik)
blokkvalim	jah/ei	Blokkvalimi kasutus (tehisõppes ei ole vajalik)
antudkood	bait	Piksliväärtus, millest kaugust või mille sagedust arvutatakse (väärtus on vajalik vaid kaugust või sagedust näitavate indeksite arvutamisel)
eraldistekasutus	jah/ei	Kas arvutada indeksid vaid eraldiste piires (vajalik vaid indeksite ja kaartide arvutamisel)
eraldistekiht	bait	Eelklassifitseerivate eraldiste kiht mustriindeksite arvutamisel
BV_kataloog	tekst	Binaarväljundi salvestamise kataloog (vajalik vaid indeksite arvutamisel binaarfaili)
asendab	bait	Asendatava tunnuse number. Kui tunnus ei ole asendustunnus, siis peab väli olema tühi või = 0.
divisor	4-baidine murdarv	Arv, millega jagatakse kahebaidise andmekihi piksliväärtused (vajalik vaid indeksite arvutamisel)
F1, F2, F3 ...	jah/ei	Argumenttunnuse kasutamine koos numbrile vastava funktsioontunnusega indeksite arvutamisel ja tehisõppel. Kasutatakse juhul, kui ftyyp ≠ 4.
AF1, AF2, AF3 ...	bait	Tunnuse aktuaalsus numbrile vastava funktsioontunnuse tehisõppel. Kasutatakse juhul, kui ftyyp ≠ 4.
F1_66 ...	jah/ei	Argumenttunnuse kasutamine koos numbrile vastava funktsioontunnuse klassiga 66 indeksite arvutamisel ja tehisõppel. Kasutatakse juhul, kui ftyyp = 4.
AF1_66 ...	bait	Tunnuse aktuaalsus numbrile vastava funktsioontunnuse antudkoodi = 66 tehisõppel. Kasutatakse juhul, kui ftyyp = 4.

Blokkvalim võimaldab kasutada rohkem külgnevaid piksleid kui lihtne juhuvalim. Blokkvalimi puhul valitakse 3×3 piksli suurusi ruute. Blokkvaliku puhul määratakse piksliruutude vahemaa nii, et kasutaja määratud valimisse sattuvate pikslite osa oleks võimalikult täpselt järgitud. Blokkvalim on vajalik külgnevaid piksleid võrdlevate tunnuste puhul, kui juhuvalimi maht on väike.

Lisa 2.3. Vajalikud väljad vaatluste andmebaasiobjektis.

Välja nimi	Vorming	Kommentaar
VID	8-baidine täisarv	Vaatluse tunnusnumber
F	4-baidine murdarv ftyyp = 2 puhul, jah/ei või bait ftyyp = 3 puhul, bait muudel juhtudel	Funktsioontunnuse väärtus (vajalik etteantud väärtusega samasuse arvutamisel, tulemuste tabelis ei ole vajalik)
progn	sama, mis väljal F	Ülejäänud õpetusandmete järgi prognoositud väärtus (vajalik vaid tehisõppel ja tabelisse prognoosimisel)
prob	bait	Sarnasus ülejäänud õpetusandmete hulgast valitud näidistega vahemikus 0...100 (vajalik vaid tehisõppel ja tabelisse prognoosimisel)
aktuaalsus	bait	Vaatluse aktuaalsus (vajalik vaid tehisõppel)
w	4-baidine murdarv	Vaatluse kaal näidisenä
x	8-baidine murdarv	Vaatluskoha lääne-ida suuna koordinaat
y	8-baidine murdarv	Vaatluskoha lõuna-põhja suuna koordinaat
alates	kuupäev, kirjutamise ja kuvamise vorm sõltub operatsioonisüsteemi suvanditest, võib olla tühi.	Vaatluse kasutatavuse algus
kuni	kuupäev, kirjutamise ja kuvamise vorm sõltub operatsioonisüsteemi suvanditest, võib olla tühi.	Vaatluse kasutatavuse lõpp
<seletava tunnuse nr>	bait	Seletavate tunnuste väärtused

Lisa 2.4. Vajalikud väljad logitabelites.

Välja nimi	Vorming	Kommentaar
ID	8-baidine täisarv	Kirje tunnusnumber
antudkood	bait	Klassikood multinominaalse tunnuse klasside ükshaaval käsitlemisel
kuupäev	kuupäev	Tulemuse salvestamise aeg
tvastavus	4-baidine murdarv	Õpetusvaatluste valimis saavutatud vaatluste ja hinnangute vastavus
kvastavus	4-baidine murdarv	Hinnangute vastavus kontrollvalimis
sumsimmax	4-baidine murdarv	Optimeeritud otsitav kogusarnasus
etn	4-baidine täisarv	Näidiseks valitud vaatluste arv
valim	4-baidine täisarv	Valimi maht
tunnusekaalud	memo	Seletavate tunnuste tehisõppel leitud kaalud salvestatud kujul: {tunnuse nr} {kaal} {tunnuse nr} jne

Lisa 2.5. Vajalikud väljad tabelis *AKIHID*.

Välja nimi	Vorming	Kommentaar
KID	2-baidine täisarv	Kihi tunnusnumber
pkylg	4-baidine murdarv	Pikslü küljepikkus meetrites
kataloog	tekst	Andmekihi failide kataloog
nominal	jah/ei	= jah, kui kihti tuleb töödelda nominaalse tunnusena
eira	bait	Andmete puudumist tähistav väärtus, mis võrdsustatakse kaardivälise alaga
alates	kuupäev, kirjutamise ja kuvamise vorm sõltub operatsioonisüsteemi suvanditest	Andmekihi kasutatavuse algus
kuni	kuupäev, kirjutamise ja kuvamise vorm sõltub operatsioonisüsteemi suvanditest	Andmekihi kasutatavuse lõpp
baite	bait	= 1 Idrisi32 baidivormingu ja = 2 Idrisi32 integer vormingu puhul. Nominaalne andmekiht võib olla vaid ühebaidine.

Lisa 2.6. Väljad tabelis *PK20t*. Kolm esimest on koha tunnuste arvutamiseks tingimata vajalikud.

Välja nimi	Vorming	Kommentaar
NR	4-baidine täisarv	Kaardilehe number
min-x	8-baidine murdarv	Kaardilehe lääneserva lääne-ida suuna koordinaat.
min-y	8-baidine murdarv	Kaardilehe lõunaserva lõuna-põhja suuna koordinaat. Ei ole vajalik, kui kasutatakse vaid ruudukujulisi kaardilehti
max-x	8-baidine murdarv	Kaardilehe idaserva lääne-ida suuna koordinaat
max-y	8-baidine murdarv	Kaardilehe põhjaserva lõuna-põhja suuna koordinaat. Ei ole vajalik, kui kasutatakse vaid ruudukujulisi kaardilehti

Lisa 2.7. Vajalikud väljad omasarnasustabelites

Välja nimi	Vorming	Kommentaar
	1 bait	Nominaalse tunnuse ühe klassi kood
	2 bait	Nominaalse tunnuse teise klassi kood
	s bait	Klasside sarnasus täisprotsentides

Lisa 2.8. Vajalikud väljad tabelis *INDEKSID*

Välja nimi	Vorming	Kommentaar
IndeksID	bait	Indeksi tunnusnumber
nimi	tekst	Indeksi nimi
nomkiht	jah/ei	= jah, kui indeks arvutatakse nominaalsest andmekihist
nomind	jah/ei	= jah, kui indeksi väärtust tuleb kasutada nominaalse tunnusena

Lisa 2.9. Mustriindeksite atribuudid süsteemis **Constud**. Kui nomkiht = *jah*, siis töödeldakse andmekihis olevaid arve nominaalse tunnuse klassikoodidena, kui nomkiht = *ei*, siis töödeldakse andmekihti arvudena, kui nomind = *jah*, siis on arvutatud indeks nominaalne, kui nomind = *ei*, siis on arvutatud indeks arvuline.

IndeksiID	Nimi	nomkiht	nomind
1	antud klassi sagedus	jah	ei
2	mood	jah	jah
3	Shannoni mitmekesisus	jah	ei
4	Lloydi ühetaolisus	jah	ei
5	dominants	jah	ei
6	klasside arv	jah	ei
7	klassikülgnevuste osa	jah	ei
8	laikude suund	jah	ei
9	klassivastavuste lähedus	jah	ei
10	eri klassi vaatluspaaride osa	jah	ei
11	kaugusega kaalutud mood	jah	jah
12	klassipiiri kaugus	jah	ei
13	kaugus etteantud klassi piirist	jah	ei
14	kaugusega kaalutud sagedus	jah	ei
15	teine mood	jah	jah
16	kolmas mood	jah	jah
101	üle keskmise väärtuste osa	ei	ei
102	keskmine	ei	ei
103	standardhälve	ei	ei
104	mediaan	ei	ei
105	Morani I 8st naaberpikslit	ei	ei
106	kaugusega kaalutud Morani I	ei	ei
107	naaberpikslite erinevus	ei	ei
108	variatsioonikoeffitsient	ei	ei
109	gradiendi suund	ei	ei
110	serva ja keskosaKerinevus	ei	ei
111	miinimum	ei	ei
112	maksimum	ei	ei
113	ekstsessitegur	ei	ei
114	üleminekute sujuvus	ei	ei
115	erinevus keskmisest	ei	ei
116	gradiendi tugevus	ei	ei
117	kauguse pöördväärtusega kaalutud keskmine	ei	ei
118	triibulisus	ei	ei

LISA 3. MUSTRIINDEKSID SÜSTEEMIS CONSTUD

3.1. NOMINAALSEST ANDMEKIHIST ARVUTATAVAD INDEKSID

Antud klassi sagedus — kasutaja poolt valitud klassikoodi protsentuaalne osa kernelis.

Mood — kõige sagedasem piksliklass kernelis.

Shannoni mitmekesisus — klasside mitmekesisus valemi $H = -10 \sum p_i \log_2 p_i$ järgi, kus p_i — klassi i osakaal kernelis.

Lloydi ühetaolisus — $10 H / \lg(s)$, kus H — Shannoni mitmekesisus, s — klasside arv kernelis.

Dominants — $100 \sum p_i^2$, kus p — klassi suhteline sagedus.

Klasside arv — erinevate klasside arv kernelis.

Klassikülgnevuste osa — sama klassi pikslite külgnevuste protsent kõigist pikslikülgnevustest. Võrdub 100%, kui kõik pikslid kernelis kuuluvad samasse klassi.

Laikude suund — 0 tähistab vertikaalseid ühe piksli laiuseid ribasid, 90 tähistab horisontaalseid ühe piksli laiuseid ribasid. Laigupiiride puudumisel kirjutatakse faili kood 255.

Klassivastavuste lähedus — sama klassi pikslite keskohtade vahemaade pöördväärtuste summa suhe kõigi pikslipaaride vahemaade pöördväärtuste summasse protsentides.

Erinevat klassi vaatluspaaride osa — erinevat klassi pikslipaaride osa protsentides kõigi kernelis olevate pikslite kõigi paaride suhtes.

Kaugusega kaalutud mood — kauguse pöördväärtusega kaalutud mood. See piksliväärtus, mille puhul seda klassi pikslite kerneli keskmest kauguste pöördväärtuste summa on suurim. Kerneli keskmes olev piksel arvesse ei lähe.

Klassipiiri kaugus — lähima klassipiiri kaugus kerneli keskmest sisendfaili pikslites.

Kaugus etteantud klassi piirist — lähima etteantud klassi kuuluva piksli kaugus kerneli keskmest sisendfaili pikslites.

3.2. NUMBRILISEST ANDMEKIHIST ARVUTATAVAD INDEKSID

Üle keskmise väärtuste osa — üle keskmise väärtusega pikslite osakaal protsentides. Näitab piksliväärtuste jaotuse asümmeetriat.

Keskmine — aritmeetiline keskmine **kernelis** või valimis olevatest piksliväärtustest.

Standardhälve — ruutjuur piksliväärtuste dispersioonist lokaalse keskmise suhtes.

Mediaan — piksliväärtus, millest mõlemal pool on võrdne arv pikslid.

Morani I 8st naaberpikslit — 8 naaberpikslit Morani I väljendatuna protsentides. 0 tähistab maksimaalset võimalikku negatiivset **ruumilist autokorrelatsiooni**, 100 tähendab ruumilise autokorrelatsiooni puudumist ja 200 — maksimaalset võimalikku positiivset ruumilist autokorrelatsiooni (sarnased piksliväärtused külgnavad).

Kaugusega kaalutud Morani I — kauguse pöördväärtustega kaalutud Morani I . Skaala sama, mis eelmisel statistikul.

Naaberpikslite erinevus — keskmine erinevus esimese astme naaberpikslite vahel.

Variatsioonikoefitsient — standardhälbe ja keskmise suhe protsentides.

Gradiendi suund — lineaarse trendpinna kalde suund **kernelis**. 255 tähistab gradiendi puudumist, 50 tähistab suurenevaid väärtusi y-telje suunas, 150 — kahanevaid väärtusi y-telje suunas, 100 — suurenevaid väärtusi x-telje suunas, 0 ja 200 — kahanevaid väärtusi x-telje suunas.

Serva ja keskosa erinevus — poolraadiuse sees olevate pikslite keskvaartuse ja poolraadiuse ja täisraadiuse vahel olevate pikslite keskvaartuse vahe. Salvestamisel liidetakse erinevusele konstant 100 ja väärtused alla 0 ja üle 250 kärbitakse mainitud piirväärtusteks. Seega on indeks vahemikus 0...250 ning väärtus 100 tähistab keskkoha ja serva erinevuse puudumist.

Miinumum — väikseim väärtus etteantud raadiuses ja valimis.

Maksimum — suurim väärtus etteantud raadiuses ja valimis.

Ekstsessitegur —
$$E = \frac{10}{n} \sum \left(\frac{x_i - \bar{x}}{\sigma} \right)^4$$
, kus x_i — pikslit väärtus, \bar{x} — piksliväärtuste keskmine, n — pikslite arv **kernelis** ja σ — piksliväärtuste **standardhälve**. Annab

suuremad väärtused erineva heledusega alade kokkupuutekohtadesse.

Üleminekute sujuvus — külgnevate pikslite homogeensus järgmise valemi järgi:

$$S = 100 \cdot \frac{\sum_i \sum_j \frac{1}{1 + (x_i - x_j)^2}}{n},$$
 kus x_i ja x_j — nii horisontaal-, vertikaal- kui ka diagonaalsuunas külgnevate pikslite väärtused ja n — pikslite arv kernelis.

Erinevus keskmisest — kerneli keskmises oleva piksli väärtuse ja kõigi kernelis olevate pikslite väärtuste aritmeetilise keskmise erinevus.

Gradiendi tugevus — lineaarse trendpinna kaldenurk kernelis. Arvutamisel korrutatakse piksliväärtuste erinevus kernelis olevate pikslite keskvärtusest ja pikslite kaugus gradiendi suunal. Gradiendi tugevus esitatakse vahemikus 0...200. 0 tähistab gradiendi puudumist, s.t. väärtustega kaalutud pikslikauguste summa gradiendil võrdub kaalumata pikslikauguste summaga gradiendil. Väärtus 10 tähendab, et piksliväärtused muutuvad gradiendil ühe piksli kaugusele liikudes keskmiselt ühe võrra.

Kauguse pöördväärtusega kaalutud keskmine — kaugust arvestatakse kerneli keskmest, keskmine võetakse piksliväärtustest.

Triibulisus — suurim erinevus 9 piksli keskmise ja 3 ühel joonel oleva piksli keskmise väärtuse vahel neljas (põhja-lõuna, ida-lääne, loode-kagu ja kirde-edela) sihis. Seejärel leitakse keskmine kõigi kernelis olevate ja valimisse langenud pikslite ümber. Suunaliste struktuuride puudumisel triibulisus = 0.

LISA 4. KÜSIMUSTE VASTUSED JA SELGITUSED

1-1. Constud suudab **vektorkujul** pindobjektidest indekseid arvutada vaid juhul, kui vektorfailid eelnevalt Idrisi rst-vormingusse rasteriseerida.

1-2. Kaugust vooluveest ja muid taolisi tunnuseid saab kaardistada, kui on olemas rasteriseeritud vooluvete andmekiht.

2-1. Seni mitte. SQL keel on küll universaalne, kuid andmebaasiga ühenduse loomiseks on tarvis tarkvaratüübile vastavat ühenduse tagajat (*connection provider*). Oracle ühenduse tagaja ei ole Constud praegusesse versiooni lõimitud.

2-2. Seni mitte.

2-3. 35

2-4. Nominaalsest andmekihist arvutatavate indeksite ID numbrid on alla 100, numbrilisest andmekihist arvutatavatel indeksitel on ID >100.

2-5. 2, 11, 15, 16 ehk **mood** , kaugusega kaalutud mood, teine mood, kolmas mood.

2-6. *[FTUNNUSED].[andmepäring]*

2-7. Üks tabel iga nominaalse tunnuse jaoks.

2-8. Järgmistel väljadel olevatel arvudel peab olema sama mõõtühik: X ja Y vaatluste tabelis ja prognooside tabelis; *min-x*, *min-y*, *max-x* ja *max-y* tabelis *PK20t* ning väljad *[AKIHID].[pkylg]*, *[ATUNNUSED].[radius]*, *[ATUNNUSED].[sradius]*.

2-9. *[FTUNNUSED].[logiID]*.

2-10. Piksli külje pikkus ühikutes, mida andmebaasis kasutatakse.

2-11. Jah.

2-12. Kasutada saab, aga kraade töödeldakse nagu ristkoordinaate.

3-1. Negatiivsete pinnavormide pindala on suurem.

3-2. Punktdiagramm.

3-3. 574. Ülesannet on lihtne lahendada SQL päringuga, kus tingimuseks on tunnuste

6 ja 7 mittevõrdsus.

```
SELECT Count(VID) FROM VAATLUSKOHAD1 WHERE [6]<>[7];
```

3-4. 4,120 ja 1,029. Piksliväärtused varieeruvad rohkem seal, kus metsa ei ole. Ülesannet on lihtne lahendada päringutega:

```
SELECT StDev([6]) FROM VAATLUSKOHAD1 WHERE [1]=0;
SELECT StDev([6]) FROM VAATLUSKOHAD1 WHERE [1]=100;
```

3-5. 8 kohas. Tulemuse annab päring:

```
SELECT Count(VID) FROM VAATLUSKOHAD1 WHERE [9]>[8];
```

3-6. 255, järve kohas ei ole mullakaarti.

3-7. Vastus sõltub ülesande nr 40 lahendusest.

4-1. 1 ja 0.

4-2. $f_{tyyp} = 4$ ehk **multinomiaalse muutuja** üksikute klasside õpe.

4-3. Sest kaalud sobitatakse õpetusvalimile, täielikult välistatud ei ole ka vastupidine tulemus.

4-4. Ei ole ja ei peagi olema, kaalud ja aktuaalsused on erinevad näitajad.

4-5. Eelneva õppevooru tulemuste hulgas salvestatud aktuaalsusi kasutatakse järgnevas õppevoorus, salvestatud kaale õppes ei kasutata.

4-6. Vastus sõltub **Constud** õppe tulemustest, mis on osaliselt juhuslik protsess.

4-7. Ainsad selle klassi vaatlused omas eelklassis. Eelklasside kasutamisel ei saa prognoosida teise eelklassi näidiste järgi ning **Constud** õppes ei saa prognoosida ka sama vaatluse järgi.

4-8. Vaatluste järgi sajab kõige vähem Sõrves. Prognoositud väärtused sõltuvad **Constud** õppe tulemustest, mis on osaliselt juhuslik protsess. Väikseim sademete hulk enamasti ei ole prognoositud Sõrve jaama, sest **Constud** õppes ei saa Sõrve andmeid Sõrve jaama kohta prognoosi arvutamisel kasutada.

4-9. Vastus sõltub **Constud** õppe tulemustest, mis on osaliselt juhuslik protsess.

5-1. Ei muuda. **Logitabeli** sirvimise aken toimetab mälujuhvis oleva tabeliga.

5-2. Vastus sõltub ülesannete 90-93 tulemustest.

5-3. Vastus sõltub ülesande 99 tulemusest.

5-4. Ei muutu.

5-5. Ei muutu.

5-6. Kui õppes kasutati valimit õpetusvaatlustest.

5-7. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess.

5-8. Taimkatte õppes kasutati põhikaardi põhialasid **eelklassifikaatorina**. Igas eelklassis on eraldi näidised.

5-9. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess.

6-1. Nullist üheni.

6-2. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess. Enamasti saadakse tulemus, et kolmas laps tabelis on tõenäoliselt nakatunud, teised mitte.

6-3. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess.

6-4. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess.

6-5. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess. Vastuse saab päringutega:

```

SELECT Avg(prob) FROM Taimkatte_test WHERE [F]=[progn];
SELECT Avg(prob) FROM Taimkatte_test WHERE [F]<>[progn] and
[progn]<>255;
    
```

6-6. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess.

6-7. Vaadeldud sademete hulk varieerub rohkem, sest Constud arvutab prognoositud väärtuse näidistel olevate väärtuste kaalutud keskmisena.

6-8. Vastus sõltub Constud õppe tulemustest, mis on osaliselt juhuslik protsess, kuid enamasti on vastuseks põld.

6-9. Prognoositavad kohad ei osalenud õppes. Prognoositavas valimis ja õppevalimis ei ole ühtegi ühist vaatlust.

- 7-1. Kasutab interpoleerimiseraldiste järgi eiratava ala eristamiseks.
- 7-2. Constud annab hoiatusteate ja kasutab ülejäänud tunnuseid.
- 7-3. Kui kaardid arvutada üksikmuutuja hinnanguna, siis jah.
- 7-4. Need kaardid ei tohiks erineda.
- 7-5. Üksuste arvu piirab eelkõige väljal *[FTUNNUSED].[arvutada]* sisse lülitatud üksuste arv.
- 7-6. Need kaardid ei tohiks erineda.
- 7-7. Üldjuhul on erinevad ja sarnasused näidistega on keskmiselt suuremad, sest need on kõige sarnasema üksuse näidised. Etteantud üksus on vaid paiguti kõige sarnasem üksus.
- 7-8. Rasterfail peab olema riskülik, aga kaardistatud ala piirid võivad olla määratud suvalise maskiga.
- 7-9. Sest kasutatud andmekihtide failide nimes on sama number.