

Mitmemõõtmeliste andmete visualiseerimine isoleeritud majandusruumis, kasutades monotoonsete süsteemide konformismiskaalat:

Uurimus Hiiumaa näitel.

*Karin Juurikas
Tallinna Tehnikaülikool
Kopli 101, 11712 Tallinn
karin@tv.ttu.ee, tel +372-6204111*

*Ants Torim
Tallinna Tehnikaülikool
Raja 15, 12618 Tallinn
torim@staff.ttu.ee, tel. +372-6202306*

*Leo Võhandu
Tallinna Tehnikaülikool
Raja 15, 12618 Tallinn
leov@staff.ttu.ee, tel. +372-6202304*

Abstrakt

Suurte tunnuste hulk teeb sotsiaalmajanduslikus uuringus standartsete statistiliste meetodite kasutamise äärmiselt raskeks. Oma artiklis kasutame regionaalmajanduslike andmete analüüsiks uutset lähenemist – monotoonsete süsteemide konformismiskaalat. See meetod põhineb igale andmestiku elemendile kaalu leidmisel, mis vastab elemendi tüüpilisusele andmestiku jaoks. Kuna sarnased elemendid omavad sarnaseid kaalusid, siis on võimalik leida ka sarnaste asulate klastrid, saades nõnda asulate sotsiaalmajandusliku klassifikatsiooni. Uurimisel võtame väikese suhteliselt isoleeritud majanduslikku tähtsust omava territooriumi, omaette maakonnana toimiva Läänemere saare – Hiiumaa. Monotoonsete tehnikate konformismiskaala kasutamine Hiiumaa puhul osutus edukaks. Iga etapp kaalumiskriteeriumide kasutamisel andis tööle uue dimensiooni, tuues välja erinevaid nüansse.

Sisulises plaanis ehk sotsiaalmajandusliku arengu seisukohast tulid tuumikutena välja Käina, Emmaste ja nende ümbruse asulad ning Kõrgessaare. Samuti eristus hulk väikeasulaid, mida võib tänase seisuga pidada potentsiaalseteks, kuid valdavalt lokaalse tähtsusega arenejateks. Maakonnakeskuse Kärkla võimalik positsioon asustushierarhias on realiseeritud.

Kaalufunktsioonide kooskasutamine andis asustumustri kohta ka adekvaatse graafilise ja kaardipildi, kus hoomatav asustumustri ülesehitus: hierarhiline struktuur, funktsionaalsed ja paikkondlikud sidemed.

Töö põhitulemuseks ongi uue efektiivse analüüsimeetodi kasutamine regionaalökonomikas ja majandusgeograafias.

Edaspidi plaanime analüüsi lisada ka teisi Eesti maakondi.

1. Sissejuhatus

1.1 Üldiselt.

Regionaalmajanduslike andmete analüüsiks on mitmeid tüüplahendusi, nagu GIS [Koperski, Han, 1998] ja statistilised meetodid [Kemeny, Snell, 1972], aga selles artiklis tutvustame me mõnevõrra uudsemat lähenemist. Oma analüüsis keskendume me hoomatavat majanduslikku tähtsust omavale suhteliselt isoleeritud territooriumile, milleks on üks Läänemere saartest - omaette maakonnana toimiv Hiiumaa. Saarelise paiknemise tõttu on Hiiumaa¹ majanduslikult ja sotsiaalselt isoleeritud territoorium 184 erineva suurusega asulaga. Selle väikese territooriumi asulate iseloomustamiseks kasutame 226 majandus- ja sotsiaalset näitajat. Suur tunnuste hulk teeb standardsete statistiliste meetodite (nagu näiteks regressioonanalüüs) kasutamise äärmiselt raskeks.

Oma uurimistöös vaatleme me võimalust kasutada taoliste andmete analüüsiks monotoonsete süsteemide meetodeid [Mullat, 1976]. Nimetatud meetodid on küll vähetuntud, kuid neid on edukalt kasutatud erinevate andmeanalüüsi probleemide lahendamiseks alates maksimaalsete klikkide leidmisest graafis ja lõpetades masinõppega [Kuusik, Lind 2000, 2003, 2004, Võhandu, 2004]. Antud artiklis keskendume me monotoonsete süsteemide konformismiskaala kasutamisele. See meetod põhineb igale andmestiku elemendile kaalu leidmisel, mis intuiitselt vastab elemendi tüüpilisusele andmestiku jaoks. Kuna sarnased elemendid omavad sarnaseid kaalusid, siis on võimalik leida ka sarnaste asulate klastrid, saades nõnda Hiiumaa asulate regionaalmajandusliku klassifikatsiooni.

Töö põhitulemuseks on uue efektiivse analüüsimeetodi kasutamine majandusgeograafias. Loodame antud uuringu baasil saada tulemusi, mis pole saavutatavad tavalise GIS või statistilise analüüsiga.

1.2 Ülevaade analüüsi metoodikast.

Monotoonsete süsteemide teooria on töötatud välja TTÜ Informaatikainstituudis ja seda on laialt kasutatud andmetabeli sisemise korrastuse leidmiseks [Mullat 1976, Võhandu, Võhandu, 2003]. Siin anname ülevaate väikesest osast teoriast, mida me antud artiklis esitatud tulemuste saavutamiseks kasutasime. Antud juhul kasutame analüüsiks kahte monotoonsete süsteemide teorias laialt kasutatavat funktsiooni, mis annavad andmetabeli reale (meie tabelis asulale) näitajate esinemissageduste alusel kaalu. Kaal iseloomustab rea tüüpilisust ehk konformsust terviktabeli (asustussüsteemi) jaoks.

¹ Hiiumaa is a 1000 km² 11000 inhabitants island in the Baltic Sea.

Olgu meil antud täisarvuline andmetabel A ridade arvuga M ja veergude arvuga N . Hiiumaa andmetabelis on väärtusteks kas 1 (näitaja kehtib) või 0 (näitaja puudub), veerud vastavad näitajatele ja read asulatele. Esitame veeru i väärtuse v sagedust kujul $f(i, v)$.

Olgu meil näiteks järgmine andmetabel A koos oma sagedustega:

Tabel 1.

j / i	1	2	3	4	5	6
1	0	0	0	1	1	0
2	0	0	1	0	1	0
3	1	0	0	0	1	0
4	0	0	0	0	0	1
5	0	1	0	0	1	0
6	0	1	0	0	1	0
7	0	0	0	0	0	1
8	0	1	0	0	1	0
$f(i, 0)$	7	5	7	7	2	6
$f(i, 1)$	1	3	1	1	6	2

Definitsioon: Igale andmetabeli reale j seab kaalufunktsioon π_{01} vastavusse kaalu :

$$\pi_{01}(j) = \sum_{i=0..N} f(i, A_{ij})$$

Meie näites:

Tabel 2.

j / i	1	2	3	4	5	6	$\pi_{01}(j)$
1	0	0	0	1	1	0	$7+5+7+1+6+6 = 32$
2	0	0	1	0	1	0	$7+5+1+7+6+6 = 32$
3	1	0	0	0	1	0	$1+5+7+7+6+6 = 32$
4	0	0	0	0	0	1	$7+5+7+7+2+2 = 30$
5	0	1	0	0	1	0	$7+3+7+7+6+6 = 36$
6	0	1	0	0	1	0	$7+3+7+7+6+6 = 36$
7	0	0	0	0	0	1	$7+5+7+7+2+2 = 30$
8	0	1	0	0	1	0	$7+3+7+7+6+6 = 36$
$f(i, 0)$	7	5	7	7	2	6	
$f(i, 1)$	1	3	1	1	6	2	

Kuna suurem kaal tuleneb sagedasemate väärtuste olemisest, siis võib seda mõista kui rea konformismi (tüüpilisuse) hinnangut antud andmetabeli jaoks.

Andmetabeli sorteerimine kaalu järgi annab andmetabelile teatud korrastuse - konformismiskaala:

Tabel 3.

j / i	1	2	3	4	5	6	$\pi_{01}(j)$
5	0	1	0	0	1	0	36
6	0	1	0	0	1	0	36
8	0	1	0	0	1	0	36
1	0	0	0	1	1	0	32
2	0	0	1	0	1	0	32
3	1	0	0	0	1	0	32
7	0	0	0	0	0	1	30
4	0	0	0	0	0	1	30
$f(i, 0)$	7	5	7	7	2	6	
$f(i, 1)$	1	3	1	1	6	2	

Analoogselt võib leida kaalud ka veergudele leides sagedused üle ridade.

π_{01} kaalufunktsiooni kasutamisel hõredates andmetabelites -see tähendab tabelites, mille enamik väärtuseid on nullid - kalduvad suurimat kaalu saama read, mis sisaldavad võimalikult palju nulle. Hiiumaa näites siis asulad, mille kohta kehtivad võimalikult vähesed näitajad. Ühest küljest näitab see selliste ridade "tüüpilisust" hõreda tabeli jaoks. Samas on tihti kasulik ka kaalufunktsioon, kus me arvestame ainult näitajate omamist ega lase näitaja puudumisel kaalu suurendada.

Definitsioon: Igale andmetabeli reale j seab kaalufunktsioon π_j vastavusse kaalu :

$$\pi_j(j) = \sum_{i=0..N} g(i, A_{ij}), \text{ kus}$$

$$g(i, A_{ij}) = f(i, A_{ij}), \text{ kui } A_{ij} > 0$$

$$g(i, A_{ij}) = 0, \text{ kui } A_{ij} = 0$$

Meie näites:

Tabel 4.

j / i	1	2	3	4	5	6	$\pi_{01}(j)$	$\pi_1(j)$
1	0	0	0	1	1	0	32	0+0+0+1+6+0 = 7
2	0	0	1	0	1	0	32	0+0+1+0+6+0 = 7
3	1	0	0	0	1	0	32	1+0+0+0+6+0 = 7
4	0	0	0	0	0	1	30	0+0+0+0+0+2 = 2
5	0	1	0	0	1	0	36	0+3+0+0+6+0 = 9
6	0	1	0	0	1	0	36	0+3+0+0+6+0 = 9
7	0	0	0	0	0	1	30	0+0+0+0+0+2 = 2
8	0	1	0	0	1	0	36	0+3+0+0+6+0 = 9
$f(i, 0)$	7	5	7	7	2	6		
$f(i, 1)$	1	3	1	1	6	2		

Antud näites on järjestus mõlema kaalufunktsiooni järgi sama, kuid see ei pruugi alati nii olla:

Tabel 5.

j / i	1	2	3	4	5	6	$\pi_{01}(j)$	$\pi_1(j)$
5	0	1	0	0	1	0	36	9
6	0	1	0	0	1	0	36	9
8	0	1	0	0	1	0	36	9
1	0	0	0	1	1	0	32	7
2	0	0	1	0	1	0	32	7
3	1	0	0	0	1	0	32	7
7	0	0	0	0	0	1	30	2
4	0	0	0	0	0	1	30	2
$f(i, 0)$	7	5	7	7	2	6		
$f(i, 1)$	1	3	1	1	6	2		

Kaalufunktsiooni π_1 võib näitaja olemasolu mõttes mõista tõmbejõudusid esitavana - mingi näitaja olemasolu saab näitaja puudumisega võrreldes asula kaalu ainult suurendada. Kaalufunktsioon π_{01} esitab tõmbe- ja tõukejõude - mingi näitaja olemasolu suurendab kaalu võrreldes näitaja puudumisega ainult juhul kui enamikul asulatel on see näitaja olemas.

1.3 Andmed.

Andmemassiivi koondatavad andmed on asulad (184) ja neid iseloomustavad näitajad (226).

Hiiumaa 184 asulat jaotuvad selgelt kahte gruppi: monofunktsionaalsed ja multifunktsionaalsed ehk suuremad (kesksemad) asulad. Enamikus (3/4) külades vaid elatakse, suurem osa tegevustest paikneb väikeses osas asulatest. Mida spetsiifilisemad on tegevused, seda suuremates ja multifunktsionaalsemates asulates need paiknevad. Seega spetsiifiliste tegevuste eelduseks on teiste tegevuste olemasolu antud kohas.

Eesti Statistikaameti (ESA) järgi jagunevad meie uurimistöös kasutatavad näitajad järgmiselt:

- *rahvastikunäitajad,*
- *ettevõtete majandusnäitajad,*
- *Era-investeeringute näitajad*

Kohalike ettevõtete majandusnäitajad saime Äriregistrist Krediidinfo vahendusel; elanike täpse statistika omavalitsustest.

Ettevõtete majandusnäitajate dünaamika on arvestatud andmekogumisperioodile eelnenud kolme aasta kohta.

Uuringus on kasutatud ka tervet hulka nominaalseid väärtusi, mida ei kajasta rahvastiku- ja majandusnäitajad, suur osa neist on saadud empiiriliste uuringute käigus: intervjuude ja vaatlusandmete põhjal, so. *mingi tegevuse/funktsiooni või väärtuse olemaolu asulas.*

Kõik meie poolt kogutud ja fikseeritud andmed on asula täpsusega

2. Andmemassiivi koostamine

Kuna tegemist on siin monotoonsete süsteemide jaotuskaalaga, kanname tabelisse täpsete arvandmete asemel näitajatena arvulisi vahemikke ehk intervale, tähistades need ühega. Kui asulas mingi tegevus/ funktsioon toimub või esineb teatav näitaja, fikseerime selle

tabelis samuti number ühega. Kui kohas antud tegevus/näitaja puudub, jääb tabeliväärtuseks automaatselt null.

Fragment terviktabelist (nullidele vastavad tühjad lahtrid):

Tabel 6.

	lapsi10	lapsi10-50	lapsi 50-100	lapsi üle 100	late ololu
	L1	L2	L3	L4	L0
KÄRDLA LINN				1	1
EMMASTE KÜLA			1		1
HALDI KÜLA	1				1
HALDREKA KÜLA					
HARJU KÜLA		1			1
HINDU KÜLA		1			1
HÄRMA KÜLA					
JAUSA KÜLA		1			1
KABUNA KÜLA					
KADERNA KÜLA					

3. Andmeanalüüs monotoonsete süsteemide konformismiskaalas.

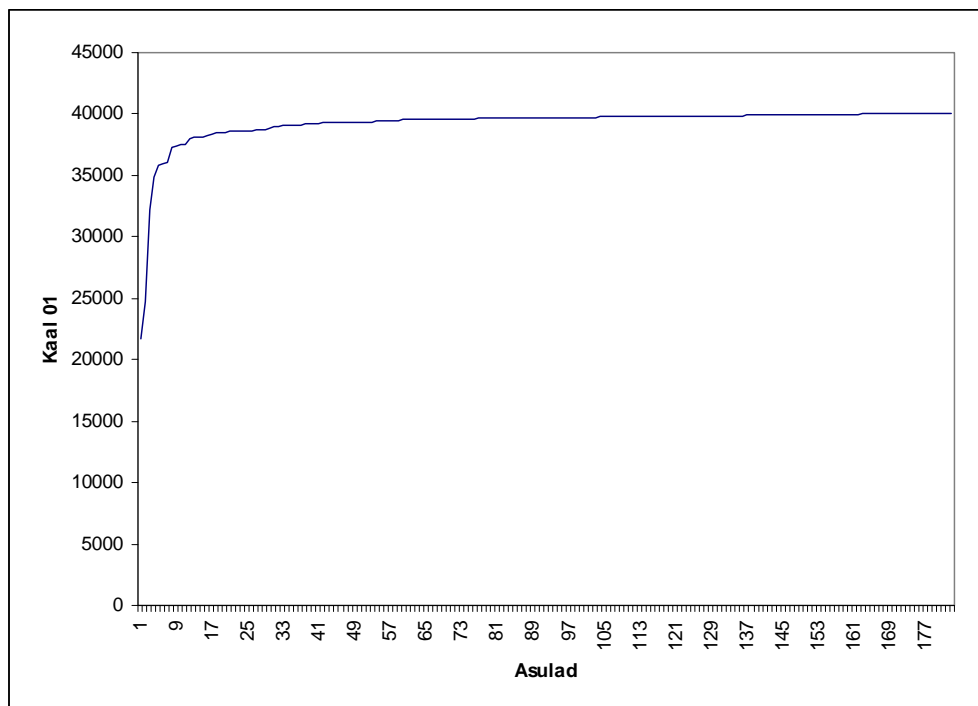
3.1. Terviksüsteemianalüüs ehk kaalufunktsioon π_{01} .

Reaalsete andmete puhul Hiiumaa asustusuuringus tuli ette, et enamikes asulates ei toimu midagi/ei täideta mingeid funktsioone, esinevad põhiliselt sotsiaalsed näitajad ning andmetabel jäi suures osas tühjaks - väärtust üks omas 4,7 % andmetabeli lahtritest.

Nullide sage esinemine saab elementide „kaalumisel“ määravaks. Seega suurema π_{01} kaalu saavad elemendid (asulad), kus on vähe näitajaid, paraku on Hiiumaal enamik selliseid külasid. Sellest lähtuvalt peaks kaalufunktsiooni π_{01} järgi Hiiumaa kõige tüüpilisematena välja tulema kohad, kus mingit majandustegevust ei toimu – inimesed elavad ja/või suvitavad seal.

Meie eeldus töötas ka reaalselt; sorteerimise tulemusena järjestati elemendid konformismiskaalas kaalu järgi, viimasteks ehk siis kõige tüüpilisemateks süsteemi elementideks (kaaluga 40029) kümme Hiiumaa eluküla: Laaritsa, Lepiku, Ulja, Aadma, Jõeküla, Ühtri, Otste, Kalgi, Pilpaküla ja Sakla.

Neist seitsmet võib pidada suuremate asulate satelliitküladeks. Analüüsitud andmetabelis on nendele kümnele külale vastavad read identsed. Kõigis kümnes on ühesugused sotsiaalsed näitajad, mis nagu me hiljem näeme on Hiiumaa külades kõige levinumad: nii lapsed tööelised kui pensionärid on esindatud, rahvaarv neis külades on 10 ja 50 vahel, nende hulgas lapsi ja pensionäre kuni kümme, ning tööelisi 10 kuni 50. Mingeid muid näitajaid need külad ei oma.



Joonis 1. Asulate jaotus terviküsteemi konformismiskaalas.

Seda juttu illustreerib ilmekalt joonis 1, mille graafikul on näha suurt hulka suure kaaluga (>36000) elemente ning ca kümnekonnal asulal on (majandusnäitajaid iseloomustavate ühede esinemise tõttu) teistest märgatavalt nõrgem kaal – kuid see-eest suurem majanduslik ja administratiivne tähtsus ning paremad sotsiaalsed näitajad (rohkem elanikke, rohkem lapsi). Seega, kõige ebatüüpilisemateks süsteemi mõistes osutusid Hiiumaa suurimad asulad eesotsas pealinn Kärdlaga (kaaluga 21 659), temale järgnesid Käina, Emmaste, Kõrgessaare jt. Kui nimetatud neli on ühtlasi ka administratiivkeskused, siis teisi suuri asulaid ühendavad sarnased sotsiaalsed näitajad.

Tabel 7. Huvitavamad asulad, sortituna kaalu π_{01} kasvamise järjekorras:

Asula	π_{01}	Olemas-olevaid tunnuseid	Selgitus
KÄRDLA LINN	21659	109	Pealinn
KÄINA ALEVIK	24733	91	Valla keskus
EMMASTE KÜLA	32221	49	Valla keskus
KÕRGESSAARE ALEVIK	34849	34	Valla keskus
KASSARI KÜLA	35795	30	
NÕMME KÜLA	35947	30	
MÄNNAMAA KÜLA	36017	27	
38 asulat
NÕMMERGA	39375	0	Tunnusteta asulad. Kuigi andmetabelis on valdavaks nullid, pole need elaniketa asulad siiski kõrgeima kaaluga.
SÜLLUSTE	39375	0	
TIHARU	39375	0	
VIITASOO KÜLA	39375	0	
LEERIMETSA	39375	0	

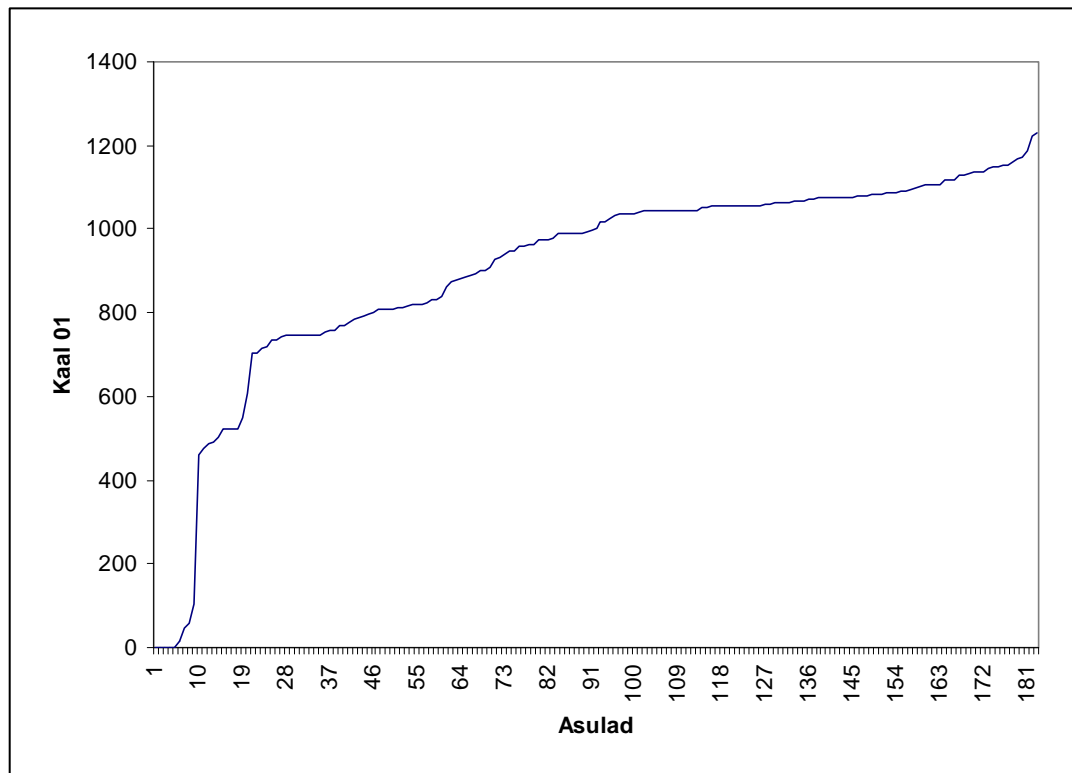
110 asulat
VIIRI KÜLA	39971	9	Olemasolevad tunnused: inimesi asulas, tööelised olemas, pensionärid olemas, lapsed olemas, pensionäre 1- 10, lapsi 1- 10, tööelisi 1-10, rahvaarv 10 - 50, suvilad
HAGASTE KÜLA	39971	9	
KUUSIKU KÜLA	40005	8	Olemasolevad tunnused: inimesi asulas, tööelised olemas, pensionärid olemas, lapsed olemas, pensionäre 1- 10, lapsi 1- 10, tööelisi 1-10, rahvaarv 10 - 50
KÖMMUSSELJA KÜLA	40005	8	
KLEEMU KÜLA	40005	8	
LAHEKÜLA KÜLA	40005	8	
MÄELTSE KÜLA	40005	8	
PÄRNSSELJA KÜLA	40005	8	
HEIGI KÜLA	40005	8	
HEISTE KÜLA	40005	8	
KIDASTE KÜLA	40005	8	
LILBI KÜLA	40005	8	
POAMA KÜLA	40005	8	
LAARTSA KÜLA	40029	8	Olemasolevad tunnused: inimesi asulas, tööelised olemas, pensionärid olemas, lapsed olemas, pensionäre 1- 10, lapsi 1- 10, tööelisi 10-50, rahvaarv 10 - 50
LEPIKU KÜLA	40029	8	
ULJA KÜLA	40029	8	
AADMA KÜLA	40029	8	
JÕEKÜLA KÜLA	40029	8	
ÜHTRI KÜLA	40029	8	
OTSTE KÜLA	40029	8	
KALGI KÜLA	40029	8	
PILPAKÜLA KÜLA	40029	8	
SAKLA KÜLA	40029	8	

Kokkuvõtteks annab järjestus π_{01} järgi pildi, kus kõige tüüpilisemateks asulateks on väikesed, 10-50 elanikuga asulad, kus majandustegevust ei toimu või väljendub see suvilate olemasolus. Väljapaistvalt ebatüüpilisteks on suured administratiivkeskused.

3.2 Reaalse tegevussüsteemianalüüs ehk kaalufunktsioon π_1 .

Meid aga huvitas nüüd asulate funktsionaalsusest (majanduslikust ja administratiivsest tegevusest) lähtuv tüüpilisus, mis ajendas meid analüüsis kasutama lisaks ka ainult ühtede kaalumist. Kaalumise toimub siin reaalse näitajate olemasolu ja hulga baasil

Nagu eeldada võiski pakkus ühtede kaalufunktsioon π_1 olulisi visuaalseid muutusi graafikutel, võrreldes π_{01} kaalufunktsiooniga., joonis 2.



Joonis 2. Asulate jaotus tegevussüsteemi konformismiskaalas.

Joonise 2 graafikul (graafiku trendijoonel) näeme asulate selget jaotumist kolme gruppi, mida me ükshaaval ka lähemalt vaatleme:

I grupp, kaaluga <400 . Tabelikaalude ja graafiku põhjal on kõige ebatüüpilisemateks asulaid, kus ei ole sotsiaalseid näitajaid ja ka muid tegevusi on vähe – nemad moodustavad siin üheksa asulaga esimese selge grupi. Siinsetes asulates pole elanikke. Suvitajad olid andmekogumisperioodil vaid Kaustes. Majandustegevusi on Lehtma sadamaasulas mitmeid (sadamategevused, piirivalve-toll, suvesöögikoht jms), kuid need on Kärddlast või mujalt juhitud ning neid ei saa seostada konkreetsete ettevõtete näitajatega kohapeal. Antud grupi küldes esineb selliseid tegevusi nagu rannapüük, mingi vaatamisväärsuse olemasolu.

II grupp, $400 < \text{kaal} < 700$. Ka teine asulate grupp on graafikul selgesti eristatav. Sellesse kuuluval üheteistkümmel asulal on nõrgemad sotsiaalsed näitajad kui Hiiumaal üldiselt. Neid ühendavateks asjaoludeks on see, et küldes elab maksimaalselt kümme inimest, lapsi pole. Kolmes neist elevad vaid pensionärid, suvitajaid on pooltes küldes. Enamik väärtusi kattub eelmise grupiga: väikesadamad, rannapüük, suvesöögikoht, vaatamisväärsus jne. Eraettevõtteid neis asulates pole.

III grupp, $700 < \text{kaal}$, on niinimetatud peagrupp, mida saab nii kaalude kui visuaalse pildi põhjal omakorda kolme alamgruppi jagada.

Kuna alamgruppid tekivad tunnuste kokkulangevuste baasil on peagrupperi moodustavate asulate sotsiaalsed näitajad väga erinevad. Ka peagrupil on oma struktuur, mida me järgnevas osas vaatluse alla võtame.

Esimeses alagrupis on erinevate sotsiaalsete näitajatega asulad, siia kuuluvad nii mõned lasteta asulad kui Hiiumaa suuruselt kolmas ja neljas asula, kuid majandustegevusi on neis vähe – enamuse lokaalse tähtsusega.

Peagrupi keskmise grupi moodustavad muuhulgas saare tähtsaimad asulad ning ühtlasi ka administratiivkeskused. Osasid sellesse gruppi kuuluvaid asulaid ühendab tugevate ettevõtte majandus ja investeringunäitajate olemasolu, teisi sarnaste sotsiaalsete näitajate olemasolu.

Kuigi visuaalselt on graafikul kaks esimest alagrupperi hoomatavad, on nende põhimõtteline vahe üsna ebaselge.

Seevastu esindab peagrupi kolmas alagrupp kindlalt Hiiumaa kõige tüüpilisemaid külasid- nemad moodustavad omaette rühma saarel kõige levinumate sotsiaalsete näitajate baasil: enamikus neist on elanike arv kümnest viiekümneni, tööelisi kõigis kümnest viiekümneni, lapsi kuni 10, pensionäre samuti kuni 10. Tegevused on Hiiumaale tüüpilised: suvitamine, rannapüük, väikesadama olemasolu jt. Kõik ettevõtmised on siin lokaalse tähtsusega. Siia gruppi kuulub ka nullide ja ühtede kooskaalumise tulemusena saadud tüüpiliste asulate grupp (Sakla, Pilpaküla jt.).

Kaalufunktsiooni π_1 abil saime teada Hiiumaa neli kõige tüüpilisemat asulat olemasolevate sotsiaalsete kui tegevus – ja majandusnäitajate arvestuses – need on Sõru , Kalana, Sakla ja Pilpaküla.

Saklat–Pilpaküla eristab Kalanast–Sõrust viimaste tegevusnäitajate suurem hulk.

Tabel 8. Huvitavamad asulad, sordituna kaalu π_1 kasvamise järjekorras:

Asula	π_1	Olemas-olevaid tunnuseid	Selgitus
NÖMMERGA	0	0	I grupp algab. Tunnusteta asulad.
SÜLLUSTE	0	0	
TIHARU	0	0	
VIITASOO KÜLA	0	0	
LEERIMETSA	0	0	
PALLI	15	1	Olemasolev tunnus: rannapüük
TOHVRI KÜLA	48	2	Olemasolevad tunnused: ujumisrand, vaatamisväärsus
LEHTMA	58	12	Olemasolevad tunnused: puidutööstusega seotud tunnused, sadamategevused, piirivalve-toll, suvesöögikoht jms
KAUSTE KÜLA	106	3	I grupp lõpeb. Olemasolevad tunnused: looduskaitseline territoorium, majutus, suvilad
LEIGRI KÜLA	461	4	II grupp algab. Olemasolevad tunnused: rahvaarv 1-10, tööelisi 1-10, tööelised olemas, inimesi asulas
JÕESUU KÜLA	474	5	Olemasolevad tunnused: rahvaarv 1-10, tööelisi 1-10, tööelised olemas, inimesi asulas, looduskaitseline territoorium.
HÜTI KÜLA	488	5	Olemasolevad tunnused: rahvaarv 1-10, tööelisi 1-10, tööelised olemas, inimesi asulas, vaatamisväärsus
SUURESADAMA KÜLA	492	10	Olemasolevad tunnused: rahvaarv 1-10, tööelisi 1-10, tööelised olemas, inimesi asulas, muuseum, rannapüük, sadam, laidudevaheline vedu, laevaremont
OGANDI KÜLA	504	4	Olemasolevad tunnused: rahvaarv 1-10, pensionäre 1-10, pensionärid olemas, inimesi asulas

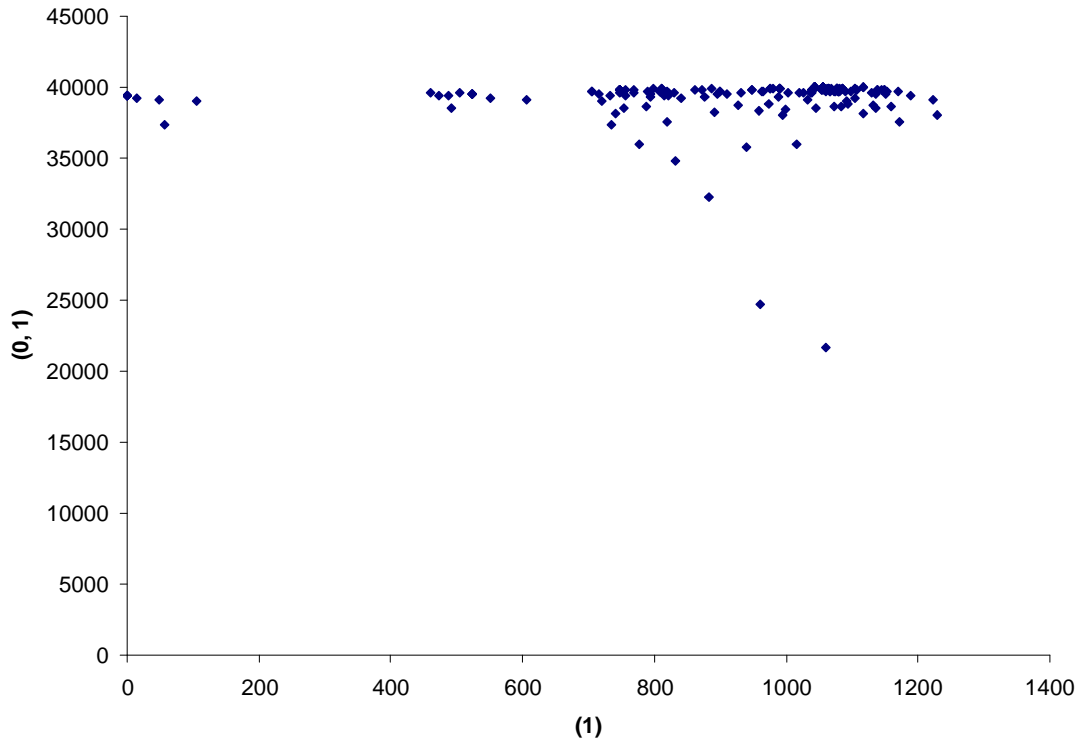
HEISTESOO KÜLA	523	5	Olemasolevad tunnused: rahvaarv 1-10, tööelisi 1-10, tööelised olemas, inimesi asulas, suvilad
NÖMME KÜLA	523	5	
TAMMISTU KÜLA	523	5	
VILIMA KÜLA	523	5	
MALVASTE KÜLA	551	7	Olemasolevad tunnused: rahvaarv 1-10, pensionäre 1-10, pensionärid olemas, inimesi asulas, majutus, suvetoitlustus
TAHKUNA KÜLA	606	8	II grupp lõpeb. Olemasolevad tunnused: rahvaarv 1-10, pensionäre 1-10, pensionärid olemas, inimesi asulas, suvilad, vaatamisväärsus, looduskaitsealine territoorium, tuulejaam
KANAPEEKSI KÜLA	705	6	III grupp algab. Olemasolevad tunnused: rahvaarv 1-10, lapsi 1-10, lapsed olemas, tööelisi 1-10, tööelised olemas, inimesi asulas.
KIIVERA KÜLA	705	6	
36 asulat
KÕRGESSAARE ALEVIK	831	34	Valla keskus
4 asulat
EMMASTE KÜLA	882	49	Valla keskus
12 asulat
KÄINA ALEVIK	960	91	Valla keskus
49 asulat
KÄRDLA LINN	1061	109	Pealinn
54 asulat
PALADE KÜLA	1173	23	Lapsi 1-10, pensionäre 1-10, tööelisi 1-10, rahvaarv 10 - 50; suvilad, raamatukogu, kool, spordihoone, muuseum, ...
SARVE KÜLA	1188	13	Lapsi 1-10, pensionäre 1-10, tööelisi 1-10, rahvaarv 10 - 50; suvilad, majutus, ujumisrand,...
KALANA KÜLA	1222	15	Lapsi 1-10, pensionäre 1-10, tööelisi 1-10, rahvaarv 10 - 50; suvilad, vaatamisväärsus, ujumisrand, rannapüük, sadam, ilmajaam, majutus ...
SÖRU KÜLA	1228	21	Lapsi 1-10, pensionäre 1-10, tööelisi 1-10, rahvaarv 10 - 50; suvilad, vaatamisväärsus, ujumisrand, rannapüük, sadam, muuseum, ...

Kokkuvõtteks võib öelda, et kaalufunktsiooniga π_l tulid kõige enam välja puhtfunktsionaalsed seosed. Asulate hierarhia tuli välja π_{0l} kaalufunktsiooniga.

3.3. Tervik- ja tegevussüsteemide ehk kaalufunktsioonide π_{0l} ja π_l omavaheline võrdlus.

Kuna kasutatavad funktsioonid tõid esile erinevaid aspekte, siis grupeerimise seisukohast annab parimaid tulemusi eelpoolkäsitletud kaalufunktsioonide tulemuste võrdlus Exceli graafikul, joonis 3. Põhimõtteliselt on tegu kahe funktsiooni koosrakendamisega.

Konformism - read



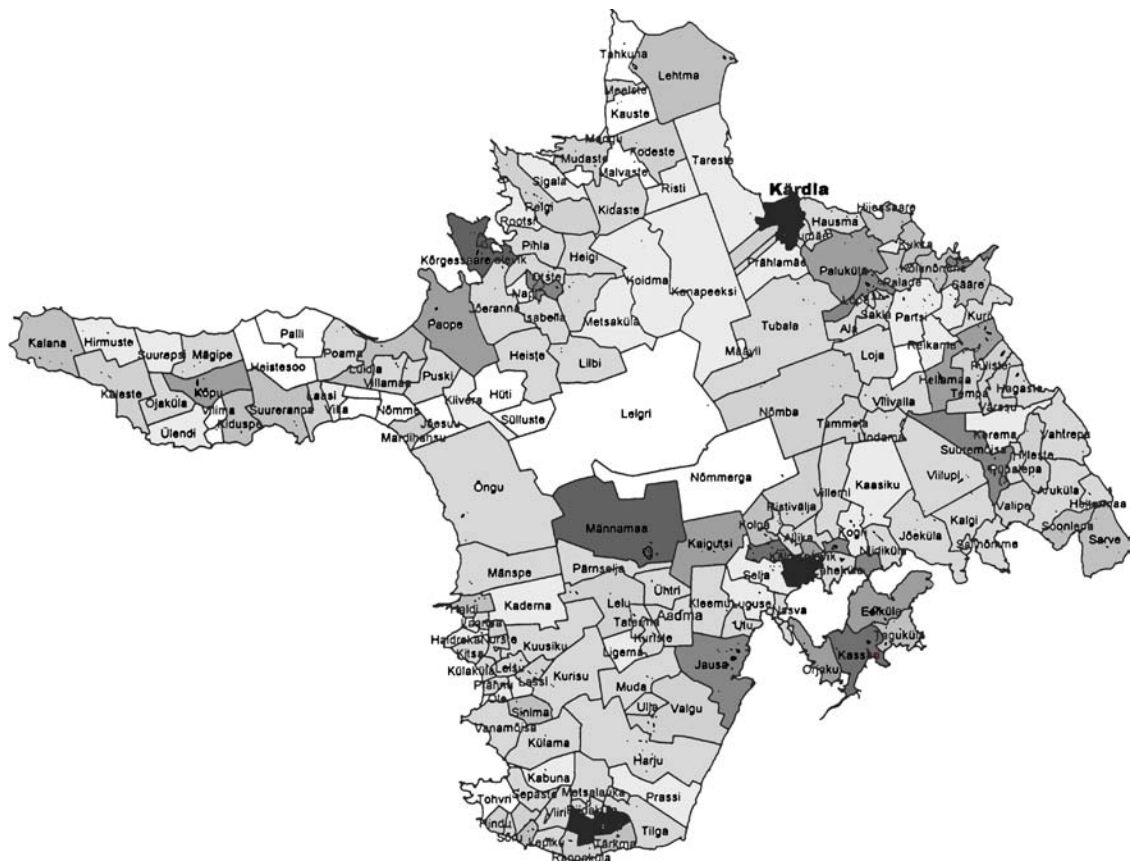
Joonis 3. Asulate jaotus π_{01} ja π_1 kaalufunktsioonide võrdluses konformismiskaalas.

Nagu näha eristuvad joonise 3 graafilises struktuuris tõmbumisskaalal π_1 kolm selget gruppi, tõukumisskaalas π_{01} eristuvad selgelt suured asulad Kärddla, Käina ja Emmaste (kaal < 35000).

Nagu juba eelmises alamosas selgus, on kaalufunktsiooni π_1 teljel eristatav "peajada", kuhu kuulub enamik elemente, ning kaks eraldiseisvat gruppi (siin) „peajadast“ vasakul. Need kaks on elaniketa (kaal < 400) ja lasteta (400 < kaal < 700) asulad. Mõlemad eelnimetatud grupid paiknevad graafikul peajadast vasakul ning nende iseloomustused on eelmises alamosas ära toodud.

Ühtede kaalufunktsiooni π_1 graafikul „peagrupid“ paiknenud suured asulad on nüüd selle all, sest nullide ja ühtede kooskaalumisel osutusid nad eranditeks. Tegevusnäitajate paljususe ehk nullide vähesuse tõttu moodustavad need nüüd 3-punktilise hajusa grupi „peajada“ all. „Peajadast“ on vähemal määral eraldumas ka teisejärgulised kesksed asulad, nagu Kõrgessaare, Kassari, Suuremõisa ning mõned majanduslikult arenevad külad (nagu Nõmmeküla).

Funktsioonide võrdlus graafikul toob välja kõikvõimalikud seotused nii „sündmuste“ toimumise (näitaja olemasolu) kui mittetoimumise asjaolust lähtuvalt ning antud tingimustes moodustavad peagrupi samade või sarnaste tunnustega punktid (asulad). Suuremad asulad on peagrupist eraldunud tänu tegevusnäitajate rohkusele (nullide vähesusele) – ning välja joonistub Hiiumaa asustussüsteemi iseloomustav muster, mis näha joonis 4 kaardipildil.



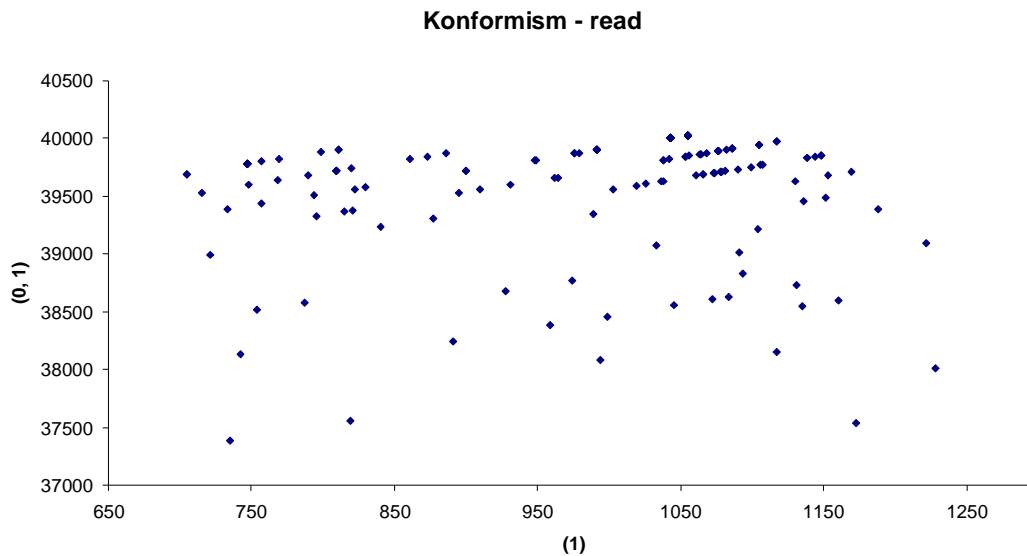
Joonis 4. Hiiumaa asustumuster konformismiskaala järgi. Tuumikalad on tumedama värvusega.

Kui kolmel Hiiumaa juhtival asulal on positsioon paigas, siis suurest grupist eraldumatel asulatel on potentsiaali edasi areneda. Välja tuli Käina ja tema ümbruse majandusarengutuumik.

4. Punktiparve sisemine struktuur

Panime tähele, et peagrupil on oma sisemine korrastatus ning punktide (asulate paiknemine graafikul allub teatavale seaduspärale, mis ka asula võimalikku arenguteed kirjeldab.

Käesolevas peatükis näitamegi peajada korrastatuse sisulist külge asustusuuringu aspektist.



Joonis 3B. Peajada

Suurendades asulate konformismiskaala pilte "peajada" koha pealt on näha diagonaaltriibud, joonis 3B. "Triipudesse" kuuluvate asulate ühiseks nimetajaks on atribuutide tõeste väärtuste arv. Kümne tõese väärtusega asulate triibu all on üheksa tõese väärtusega asulate triip, selle all kaheksa tõese väärtusega asulate triip jne. See on kaalufunktsioonide π_1 ja π_{01} matemaatiline omadus kahendtabelite korral, mille me järgnevalt tõestame.

Teoreem: Olgu meil antud suvaline kahendtabeli A rida j . Muudame andmetabelit A võttes elemendi A_{aj} , mis varem oli võrdne ühega ja asendame selle nulliga ja võttes elemendi A_{bj} , mis varem oli võrdne nulliga ja asendame selle ühega. Lihtsamalt öeldes liigutame me reas j suvalist ühest väärtust ühest veerust teise. Tähistame saadud andmetabelit A' ja rea j kaalusid selles tabelis $\pi'_{01}(j)$ ja $\pi'_1(j)$. Sel juhul:

$$\pi'_{01}(j) - \pi_{01}(j) = 2 \cdot (\pi'_1(j) - \pi_1(j))$$

Ehk ühe tõese väärtuse ümberpaigutuse korral on kaalu π_{01} muutus võrdne kahekordse kaalu π_1 muutusega.

Tõestus:

$\pi_1(j) = C + f(a, 1)$, Kus C on ühestele veergudele peale a vastavate sageduste summa

$\pi'_1(j) = C + f(b, 1)$

$(\pi'_1(j) - \pi_1(j)) = f(b, 1) - f(a, 1)$

$\pi_{01}(j) = K + f(a, 1) + f(b, 0)$, Kus K on ühestele veergudele peale a ja nullistele veergudele peale b vastavate sageduste summa

$\pi'_{01}(j) = K + f(a, 0) + f(b, 1)$

$$(\pi'_{01}(j) - \pi_1(j)) = f(b, 1) + f(a, 0) - f(a, 1) - f(b, 0)$$

Kuna igas veerus i on sagedus $f(i, 0) = M - f(i, 1)$ (M on ridade arv), siis:

$$(\pi'_{01}(j) - \pi_1(j)) = f(b, 1) + M - f(a, 1) - f(a, 1) - M + f(b, 1)$$

$$(\pi'_{01}(j) - \pi_1(j)) = 2 \cdot (f(b, 1) - f(a, 1))$$

$$(\pi'_{01}(j) - \pi_1(j)) = 2 \cdot (\pi'_1(j) - \pi_1(j))$$

qed.

Kuna igast reast on võimalik piisava arvu ühtede ümbertõstmise abil saada suvaline sama ühtede arvuga rida, siis on lihtne näha, et juhul kui read $j1$ ja $j2$ omavad sama arvu ühtesid:

$$\pi_{01}(j1) - \pi_{01}(j2) = 2 \cdot (\pi_1(j1) - \pi_1(j2))$$

Punktiridadel ehk “triipudel” on tähendus mõlemat pidi, nii vasakult-paremale kui ülevalt alla: tunnuste arv triibus kasvab, niisamuti ühiste tunnuste arv triibus: kui kõrvuti asetsevate punktide näitajad on kohati erinevad, siis punktidel ülevalt alla on täpselt samad tunnused, kuid alumised punktid omavad suuremat näitajate hulka. Samuti näitavad ülevalt-alla read vasakult paremale suunas punktide tunnuste kokkulangevuse hulka, vasakul ääres on see väikseim, paremal ääres suurim. Põhjuseks, et vasakus ääres olevatel asulatel on vähem tunnuseid ning seega ka vähem võimalikke tunnuste kokkulangevusi. Tüüpilisi, paremas ääres olevaid külasid ühendab hulk samasugused (Hiiumaal enamlevinud) näitajaid – võimalikke kokkulangevusi ongi rohkem.

Graafiliselt vasakule jäävad punktid on süsteemi jaoks ebatüüpiliste tunnustega asulad. Sellised on näiteks inimeste olemasoluga asulad, mis jäävad graafikul vasakule äärde, moodustades omaette selge grupi. Ka Hiiumaa suurimad asulad paiknevad suure grupi (punktisarve) vasakus ääres, sest nad omavad mitmeid tunnuseid, mida teistel asulatel üldjuhul pole. Näiteks, süsteemi jaoks tüüpilised asulad, sadamaasulad Kalana ja Sõru paiknevad graafikul paremas ääres üksteise all, moodustades seega rea ülevalt alla. Sõru paikneb allpool, seda tema tunnuste üldhulk on suurem kui Kalanal. Nende tunnuste ühisosa on kaksteist kokkulangevust.

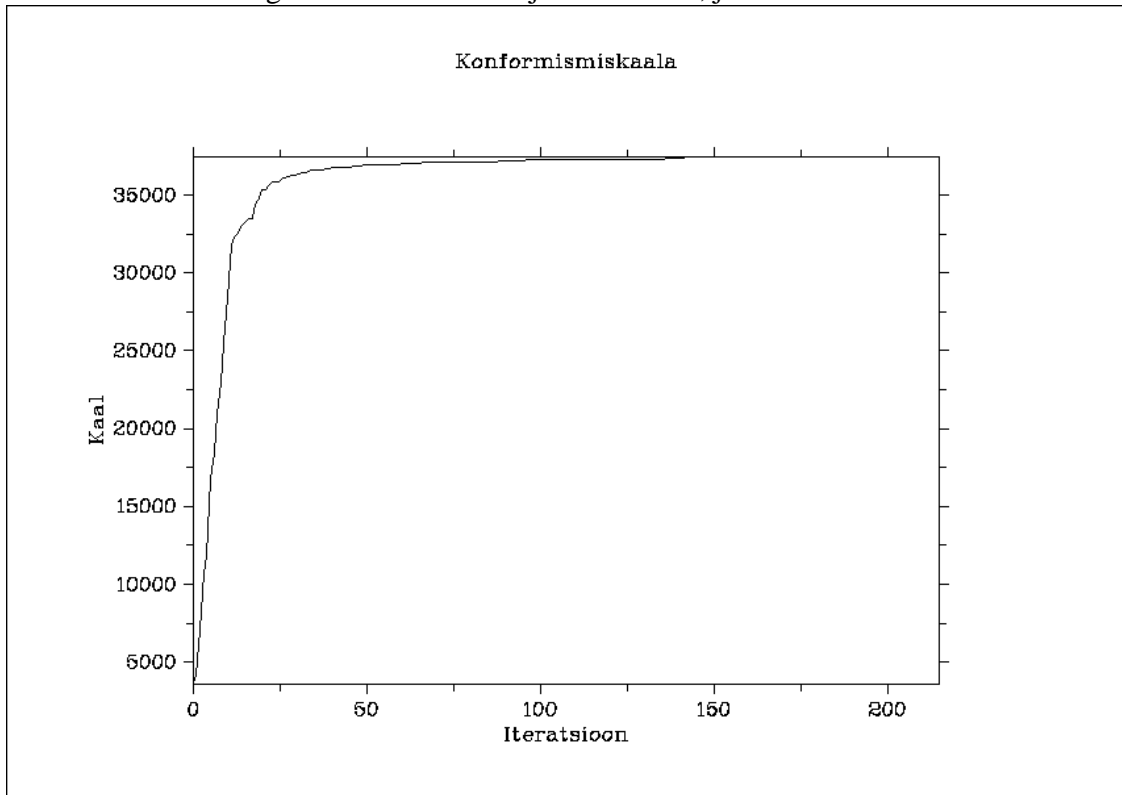
Kuna π_{01} kaalufunktsioon toob erinevalt π_1 kaalufunktsioonist esile punktidevahelised tõukejõud on seletatav ka Hiiumaale tüüpilisemate asulate grupi kokkupressitus, ning antud skaalas nagu ka joonisel 3 näha eristuvad suured teistest rohkemate tunnustega suured asulad.

5. Tunnuste jagunemine konformismiskaalas

Tunnused ehk tegevusnäitajad jagunevad omakorda ettevõtete majandusnäitajateks (näitajate intervallid, näitajate kasvu/kahanemise trend jne), lokaalseteks tegevusteks ning sotsiaalseteks näitajateks.

Hiiumaal on vähe majanduslikku tähtsust omavaid tegevusi ja nendega on samuti nagu asulatega, enamik funktsioone/näitajaid on realiseeritud kord või kaks-kolm suuremates

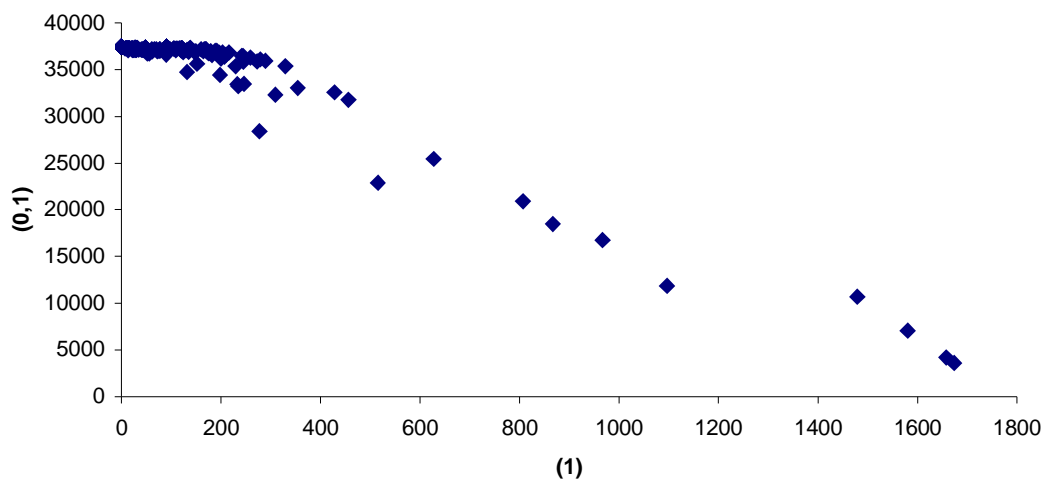
asulates või mõnes vahepeelses väikest tööstust omavas külas, kuid arvestatavat hulka külasid ühendab mingite sotsiaalsete näitajate olemaolu, joonis 5.



Joonis 5. Tegevuste jaotus terviksüsteemi konformismiskaalas. Väike osa näitajaid (sotsiaalsed näitajad) esinevad enamikus külades. Ülejäänud näitajaid esineb süsteemis väga vähe.

Tegevuste graafiku konformismiskaalas on (nii tõbumist kui tõbumist-tõukumist arvestades) välja joonistunud üks tegevuste peagrupp, sellest eristuvad kõige olulisemaid sotsiaalsete näitajate punktigrupid: nii tõbumis- kui tõukumiskaala kõige eraldunum grupp on teatavate sotsiaalsete näitajate olemaolu fakti kinnitavate näitajate grupp (tunnustega inimeste olemasolu, tööaliste olemasolu, laste olemasolu, pensionäride olemasolu asulas). Teine selgelt eraldunud grupp kajastab kõige levinumaid sotsiaalseid näitajaid (rahvaarv 10-50, tööaliste arv 10-50, pensionäre kuni 10, lapsi kuni 10) ning suvitajate olemasolu asulas. Kõik majandusnäitajad ja vähemlevinud sotsiaalsed näitajad on koondunud ühte suurde gruppi, joonis 6.

Konformism-veerud



Joonis 6. Tegevuste jaotus π_{01} ja π_1 kaalufunktsioonide võrdluses. Mõlema kaalufunktsiooni puhul tulevad välja kõige levinumate sotsiaalsete näitajate grupid.

Konformismiskaalas moodustunud suur grupp jaguneb oma sisemise struktuuri baasil samuti väiksemateks gruppideks, mis eristuvad selle järgi, mitmes asulas nad kehtivad: üks kord ehk ühes asulas kehtivate näitajate grupp, kaks korda ehk kahes asulas kehtivate näitajate grupp, kolmes, neljas jne.

π_1 kaalufunktsiooni skaalas eristuvad omakorda tüüpilisemad ja ebatüüpilisemad näitajad, mis toovad omakorda välja saare tüüpilisemad (arvustussüsteemi enda jaoks kõige ebatüüpilisemad) tegevusnäitajad, joonis 6.

Kokkuvõte

Edukas osutus monotoonsete tehnikate konformismigraafiku kasutamine majandusgeograafilises uuringus väikese poolavatud territooriumi – Hiiumaa puhul. Iga etapp kaalumiskriteeriumide kasutamisel andis tööle mingi uue dimensiooni, tuues välja erinevaid nüansse:

- (1) Esimeses etapis, kui arvutussüsteem arvestas elementide (asulate) mõõtmisel nii ühtede kui nullide sagedust eristusid kõige ebatüüpilisematena kõige enam näitajaid ja tegevusi omavad suuremad asulad. Kõige tüüpilisemateks osutusid kõige levinumate sotsiaalsete näitajatega elu- ja suvituskülad.
- (2) Kuna Hiiumaa küldes on vähe tegevusi, koosneb andmetabel arvutussüsteemi jaoks peamiselt nullidest. Seda asjaolu arvestades proovisime kasutada ka ainult ühtede kaalufunktsiooni, sellega tulid kõige puhtamalt välja funktsionaalsed (tegevuspõhised) seosed, mistõttu on kaardimudelil jälgitav isegi maanteede kulg, kuigi vastavaid andmeid pole töötlustabelisse sisestatud.
- (3) Parimaid tulemusi andis kaalufunktsioonide omavaheline võrdlus, kus eristusid Hiiumaa mõistes suured asulad, inimtühjad asulad ning madalate sotsiaalsete

näitajatega lasteta asulad. Kuna kasutatavad funktsioonid töid esile erinevaid aspekte, siis grupeerimise seisukohast annab parimaid tulemusi eelpoolkäsitletud kaalufunktsioonide tulemuste võrdlus Exceli graafikul ehk funktsioonide kooskasutamine.

(4) Sarnaselt asulatele töötas konformismiskaala ka tunnuste jaotuse puhul.

Sisulises plaanis ehk majandusarengu seisukohast tuli tuumikuna välja Käina ja selle ümbrus (Nõmmeküla, Kassari, jt.). Käina alevik on perspektiivse majandustuumiku keskmeks.

Kärdla võimalik positsioon asustushierarhias on realiseeritud. Metsade keskel ja Palade-Partsi asuala äärealal paiknev Kärdla on hälve üldlevinud asustustrendist – kunagise kalevivabriku linnakuks ehitatud asula edasine areng on võimalik vaid tsentraliseerimise poliitika võtmes või mõne olulise majanduslikku kaalu omava ettevõtte lisandumisel. Analüüsi tulemuste põhjal osutus ta nõrgemaks keskuseks, kui eeldasime.

Piirkonna majandusarengu potentsiaali mõttes on teisel kohal Emmaste. Elanike arvult on Emmaste alles viies asula Hiiumaal, kuid ta kuulub näitajate baasil Hiiumaa tippasulate kolmikusse. Tema ümbruses on mitmeid olulist majanduslikku kaalu omavate ettevõtetega asulat.

Kõrgessaare, kus andmekogumisperioodil toimus kogu asulat mõjutanud suure ettevõtte (kalavabriku) pankrot, puhul on tegemist samuti perspektiivika paigaga, kui edaspidi tehakse õigeid juhtimisotsuseid. Tänapäevase seisuga on vabrikul juba uued omanikud ning töö käib edasi.

Graafilisel pildil „peagrupid“ allapoole eemalduvaid väikeasulaid võib tänapäevase seisuga kõiki pidada potentsiaalseteks, kuid valdavalt lokaalse tähtsusega arenejateks.

Eriti huvitavaks võib osutada π_7 kaalufunktsiooniga kõige tüüpilisemateks külade - Sõru ja Kalana areng. Mõlemad külad eenduvad „peagrupid“, nii vertikaalses kui horisontaalses plaanis, tänapäevaseks on mõlemad oma arengupotentsiaali jõudnud juba mõnevõrra realiseerida Sõru omavalitsuse ning Kalana eraisikute initsiatiivil.

Kaalufunktsioonide kooskasutamine andis asustumustri kohta ka adekvaatse graafilise ja kaardipildi, kus hoomatav asustumustri ülesehitus: hierarhiline struktuur, funktsionaalsed ja paikkondlikud sidemed.

Edaspidi plaanime analüüsi lisada ka teise Läänemere saare –Saaremaa.

Kasutatud kirjandus.

1. J.G. Kemeny, J.Laurie Snell. 1972. *Mathematical Models in the Social Sciences*. The MIT Press, Cambridge, Massachusetts, and London, England. Pp. 145
2. L.Võhandu, P.Võhandu. Simple and effective methods of data handling in risk analysis. *Risk and Safety Management in Industry, Logistics, Transport and Military Service: New Solutions for the 21st Century*. Proceedings of the international scientific-educational conference. US Office of Naval Research International Field Office. Technical University of Tallinn. March 25-28, 2003 Tallinn. Pp. 37-40.
3. Mullat, I. Extremal Monotonic Systems. *Automation and Remote Control* 1976 No 5
4. Kuusik, R. Lind, G., Võhandu, L. Pattern Mining as a Clique Extracting Task. Posters. Tenth International Conference IPMU 2004 Information Processing and Management of Uncertainty on Knowledge-Based Systems. July, 4-9, 2004, Perugia, Italy, ISBN 88-87242-54-2, pp. 19-20.
5. Lind, G. Method for Data Mining - Generator of Hypotheses. *Databases and Information Systems*. Proceedings of the 4th International Baltic Workshop. Vilnius 2000. Vol. 2, pp. 304-305.
6. Kuusik, R., Lind, G. An Approach of Data Mining Using Monotone Systems. Proceedings of the Fifth International Conference on Enterprise Information Systems. Angers, France 2003. Vol. 2, pp. 482-485.
7. Kuusik, R., Lind, G. New frequency pattern algorithm for data mining. Proceedings of the 13th Turkish Symposium on Artificial Intelligence and Neural Networks. 10-11 June, 2004, Foca, Izmir, Turkey, ISBN 975-441-213-8, pp. 47-54.
8. K. Koperski, J. Han, and J. Adhikary. Mining Knowledge in Geographical Data. In *Communications of ACM*, 1998.