

Korelācijas koeficienta aizvietošana ar svērto korelācijas koeficientu ģeogrāfisko pētījumu datu analīzē

Juris Paiders

**Ģeogrāfiskajos pētījumos
izplatīta matemātisko metožu
forma ir regresijas analīze un
korelācijas vai determinācijas
koeficientu noskaidrošana
starp dažādu indikatoru
rādītājiem**

$$r = \frac{(n \times \sum x_i \times y_i) - (\sum x_i) \times (\sum y_i)}{\left(\sqrt{(n \times \sum x_i^2) - (\sum x_i)^2} \right) \times \left(\sqrt{(n \times \sum y_i^2) - (\sum y_i)^2} \right)}$$

- kur r - lineārais korelācijas koeficients;
- n - novēroto pāru skaits;
- x, y - analizējamo pazīmju pāru vērtības

regresijas koeficients (b)

$$b = \frac{\left(n \times \sum x_i \times y_i\right) - \left(\sum x_i\right) \times \left(\sum y_i\right)}{\left(n \times \sum x_i^2\right) - \left(\sum x_i\right)^2}$$

- kur a - vienādojuma brīvais loceklis;
- b - regresijas koeficients;
- r - lineārais korelācijas koeficients;
- n - novēroto pāru skaits;
- x, y - analizējamo pazīmju pāru vērtības

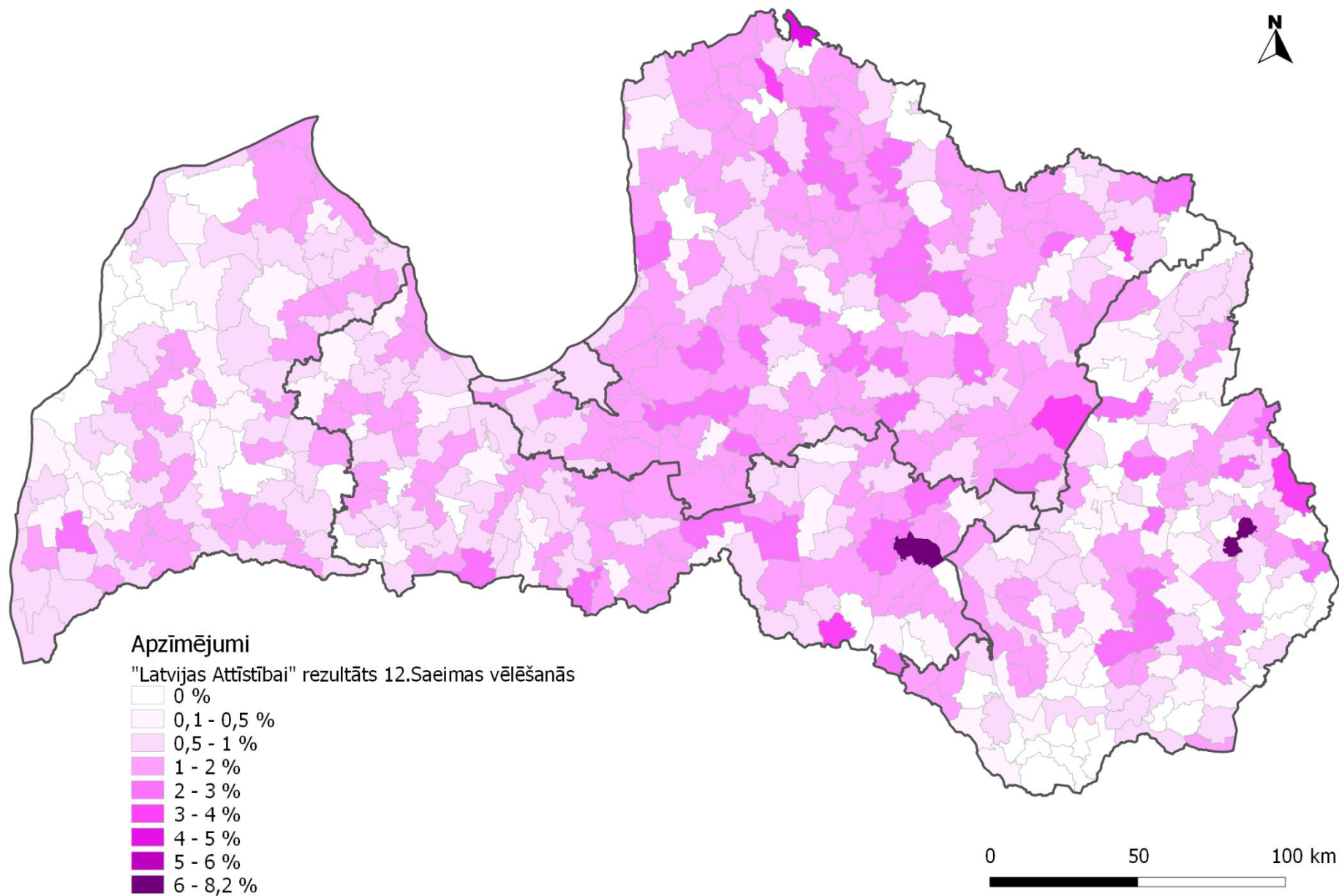
vienādojuma brīvais loceklis (a)

$$a = \frac{\sum y_i - b \times \sum x_i}{n}$$

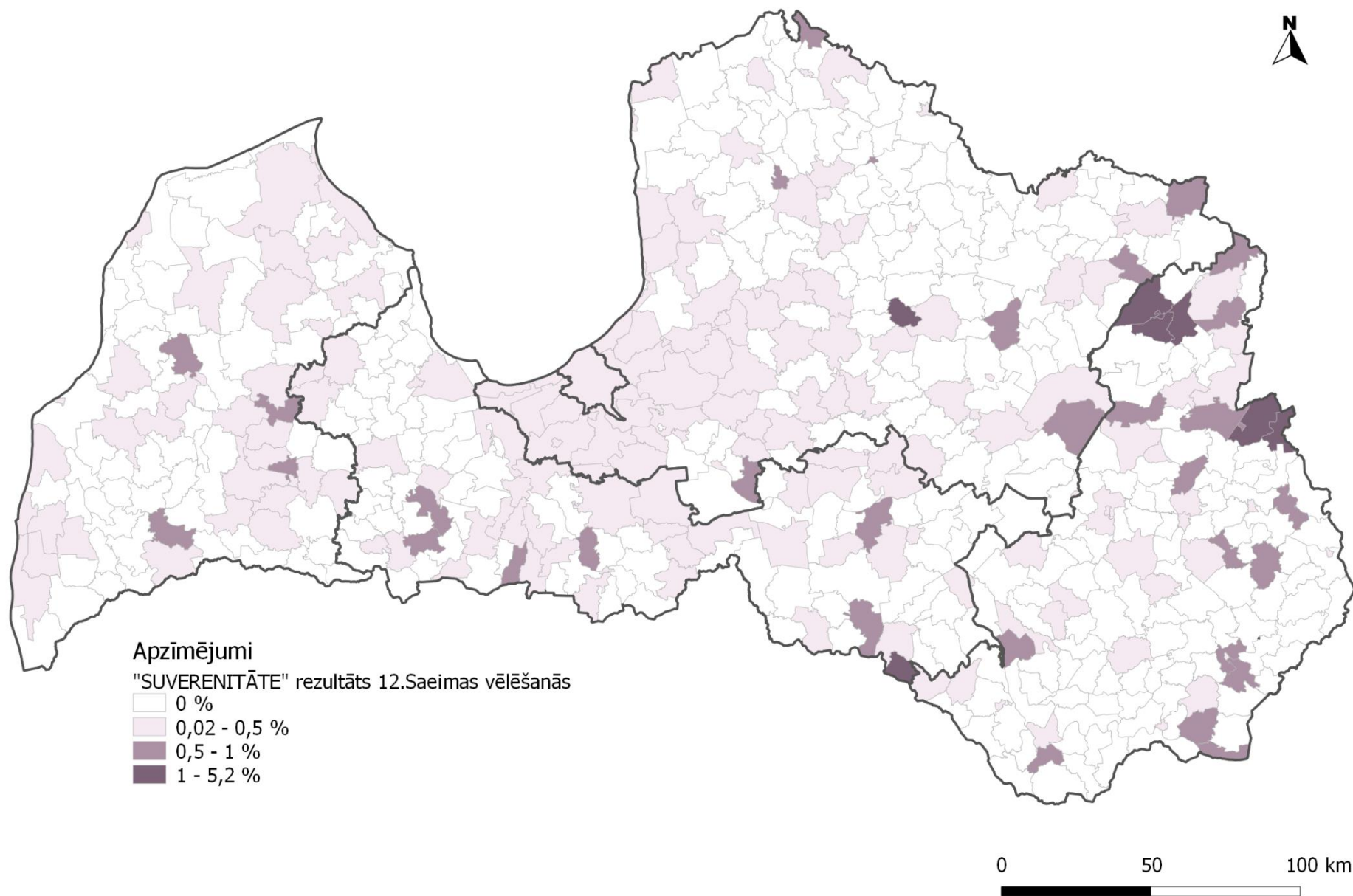
Veicot regresijas analīzi pa teritoriālajām vienībām, matemātiski korekts rezultāts būs tikai tad, ja ikvienas teritorijas ietekme un nozīme ir vienāda ar pārējām teritorijām

12. Saeimas vēlēšanu rezultātus noteicošie faktori visā Latvijā

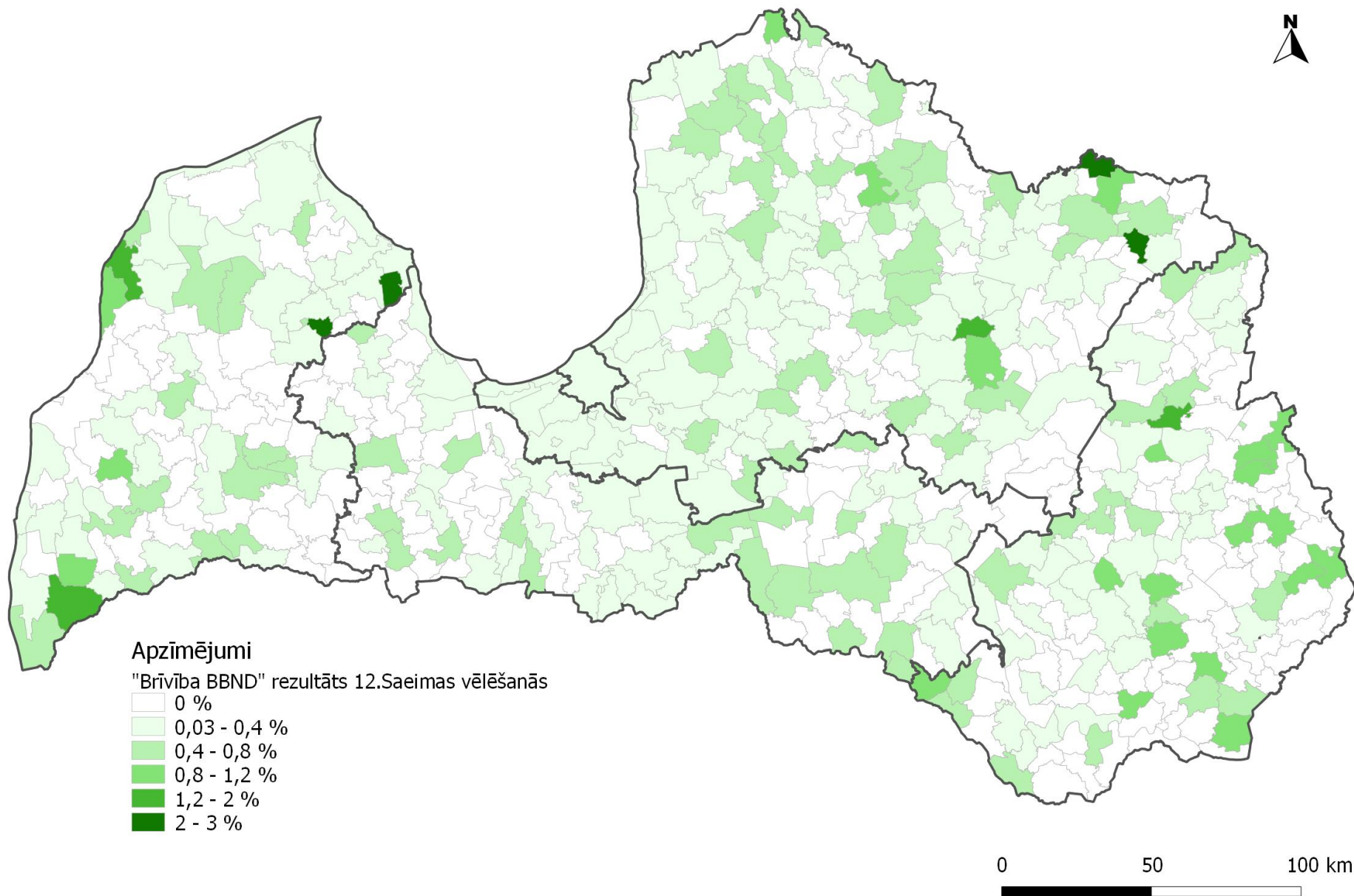




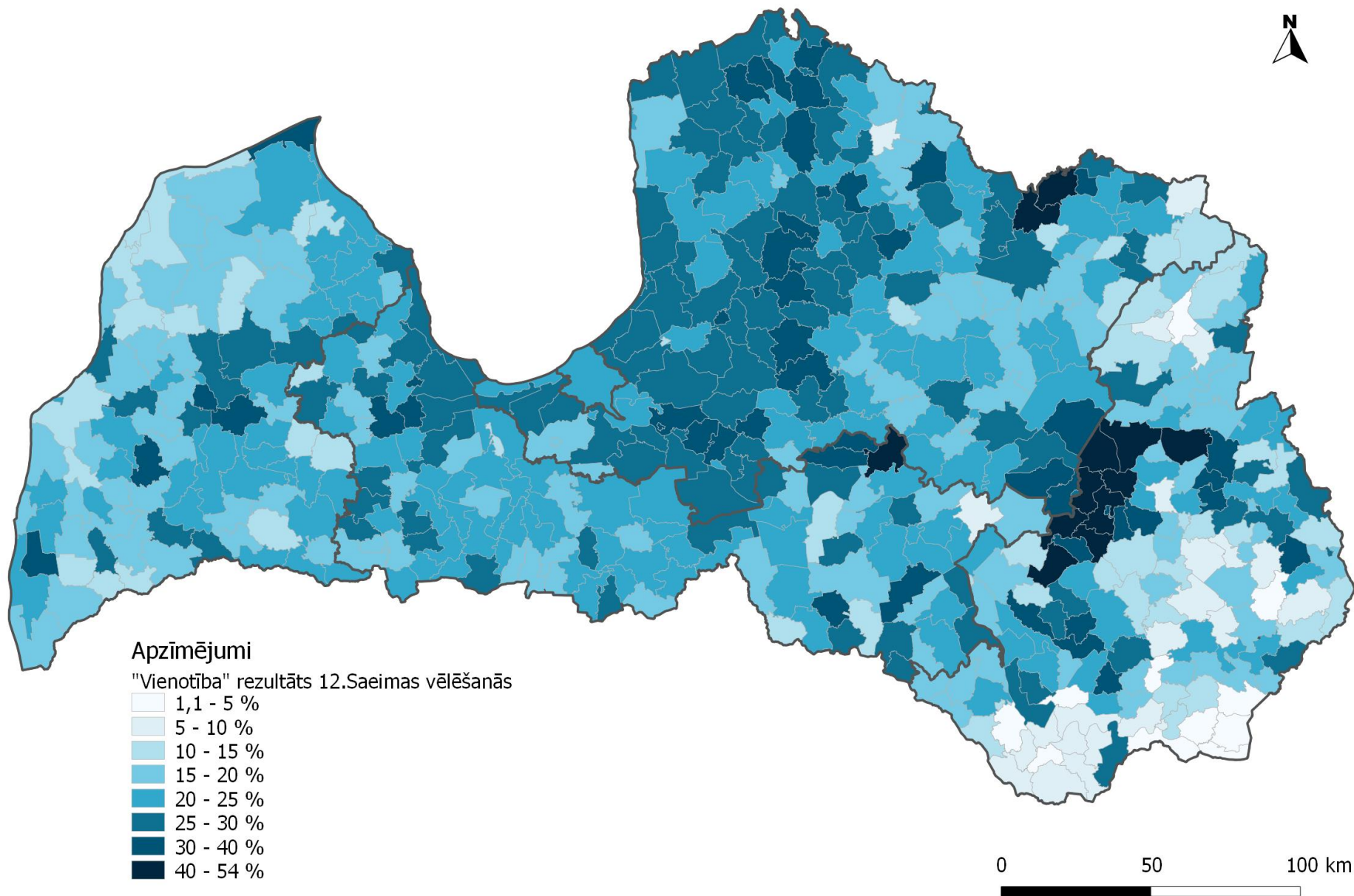
Avots: Jānis Paiders, izmantojot CVK (2014)



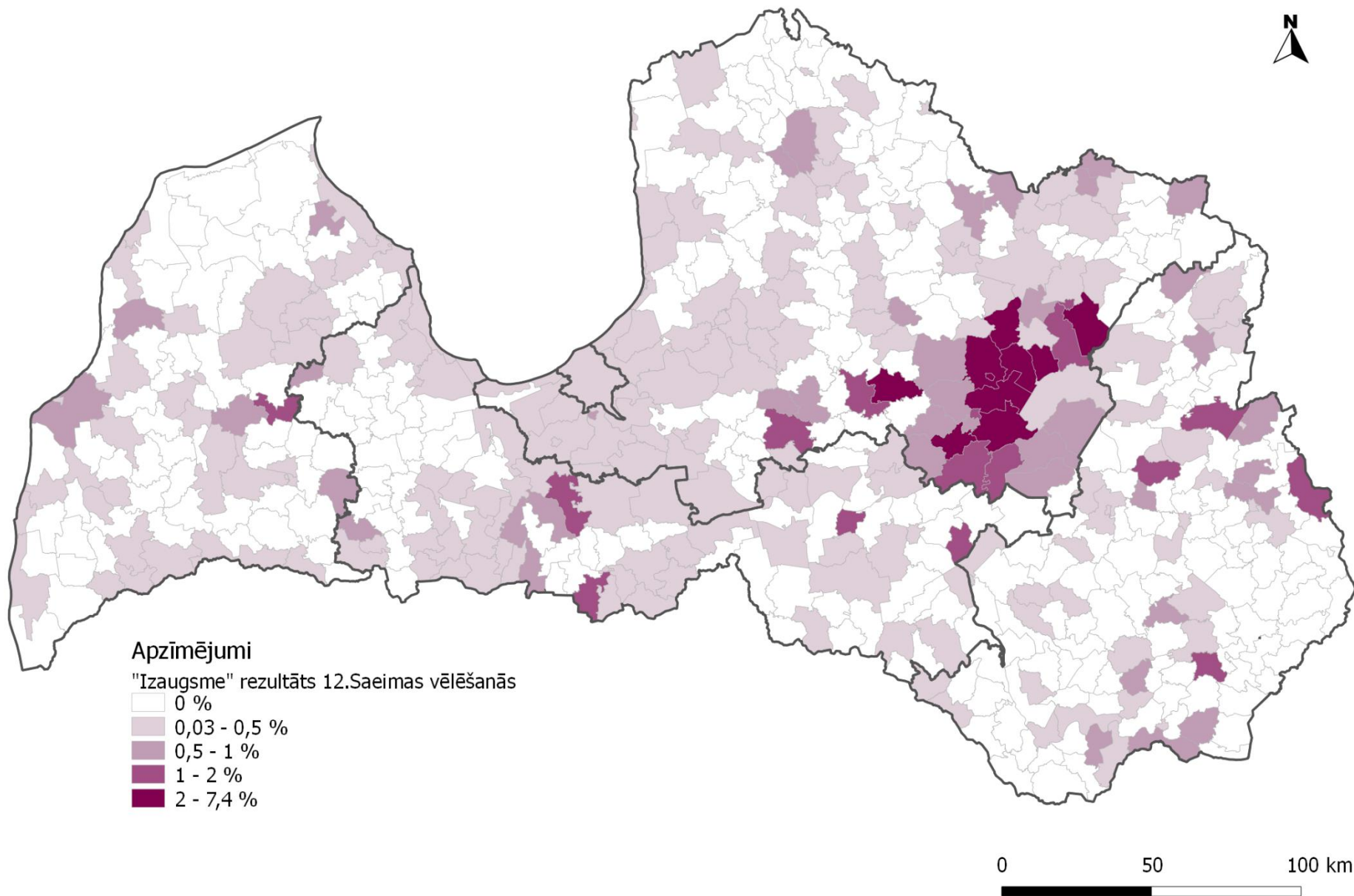
Avots: Jānis Paiders, izmantojot CVK (2014)



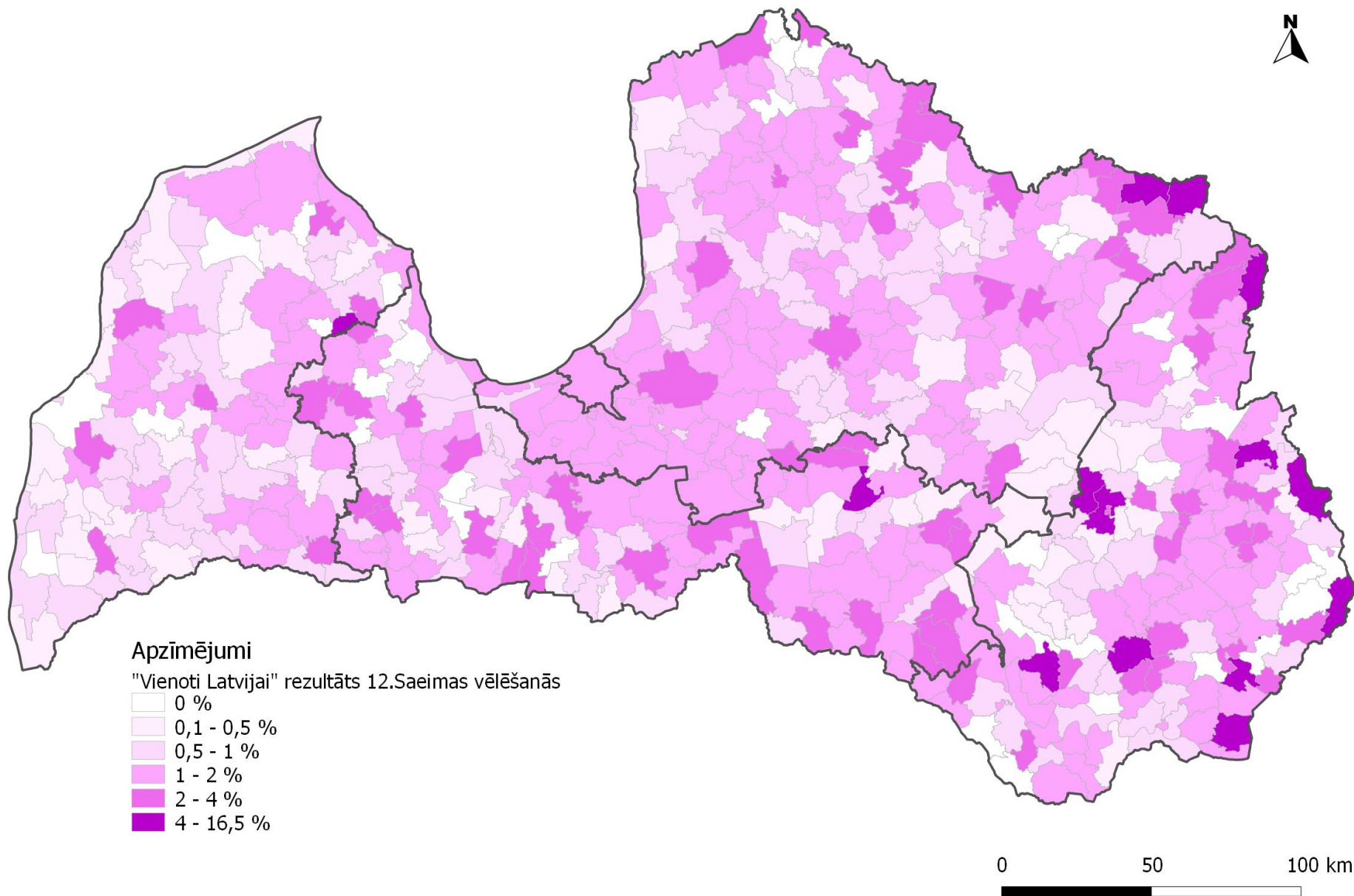
Avots: Jānis Paiders, izmantojot CVK (2014)



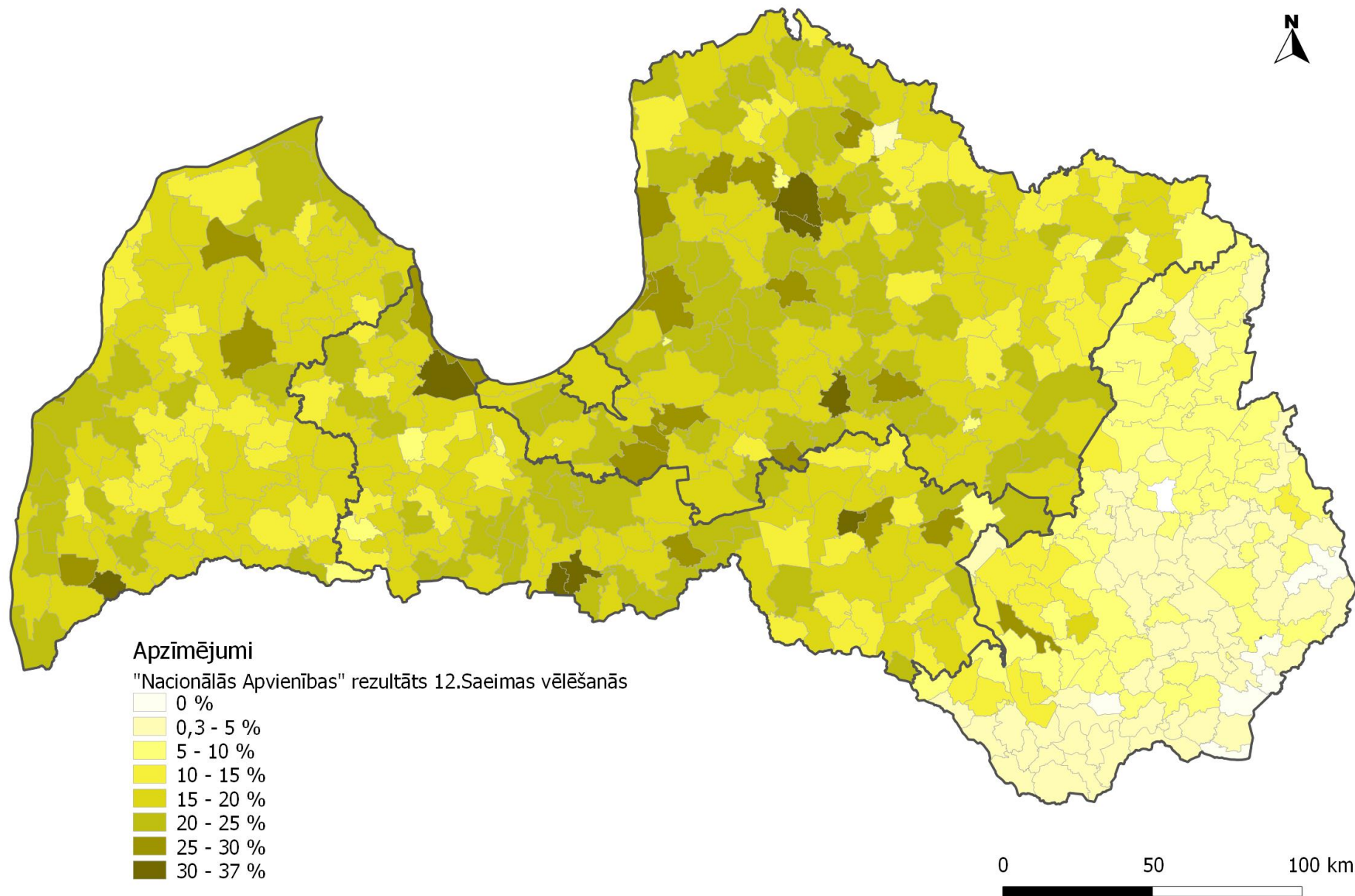
Avots: Jānis Paiders, izmantojot CVK (2014)



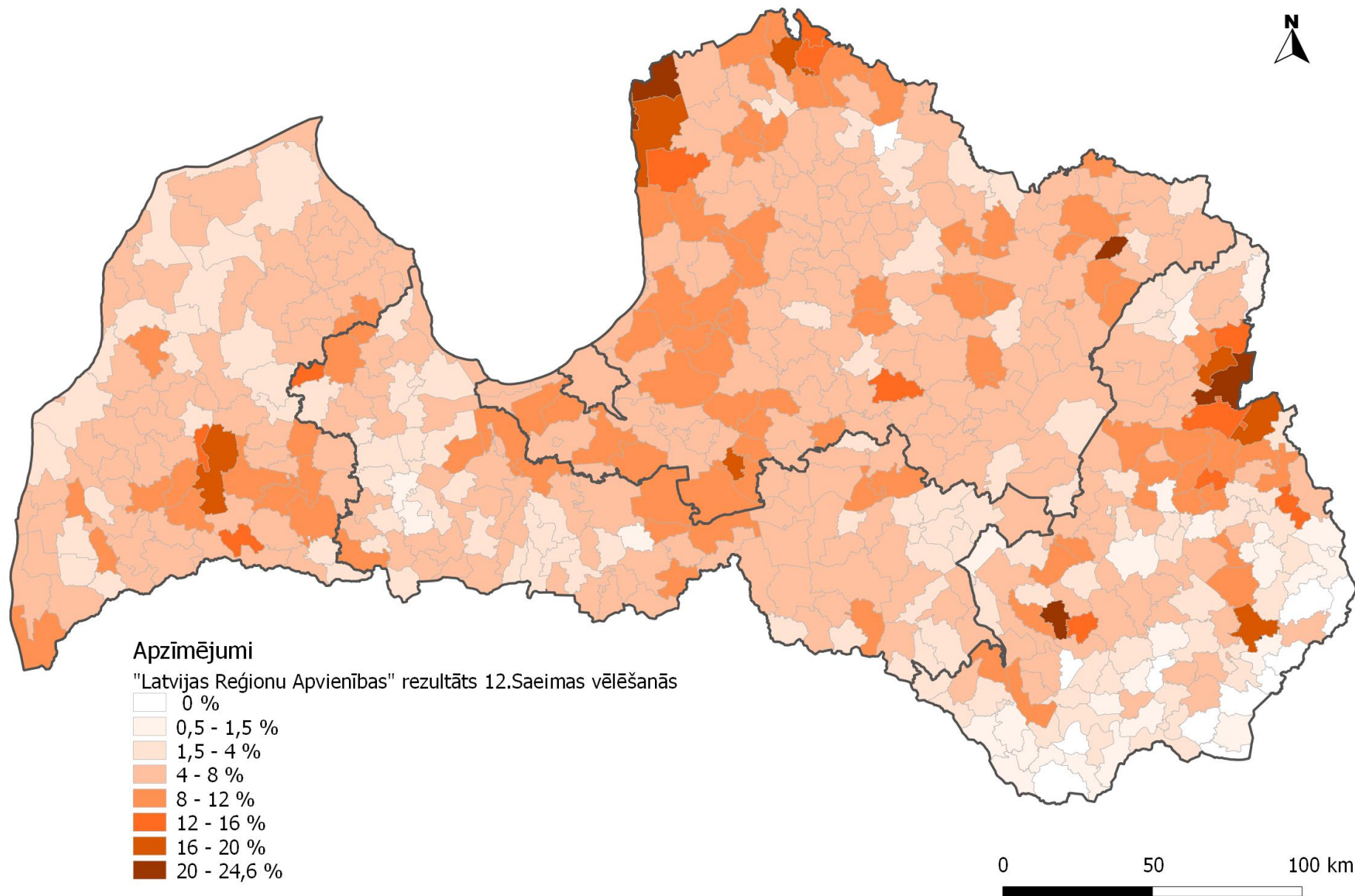
Avots: Jānis Paiders, izmantojot CVK (2014)



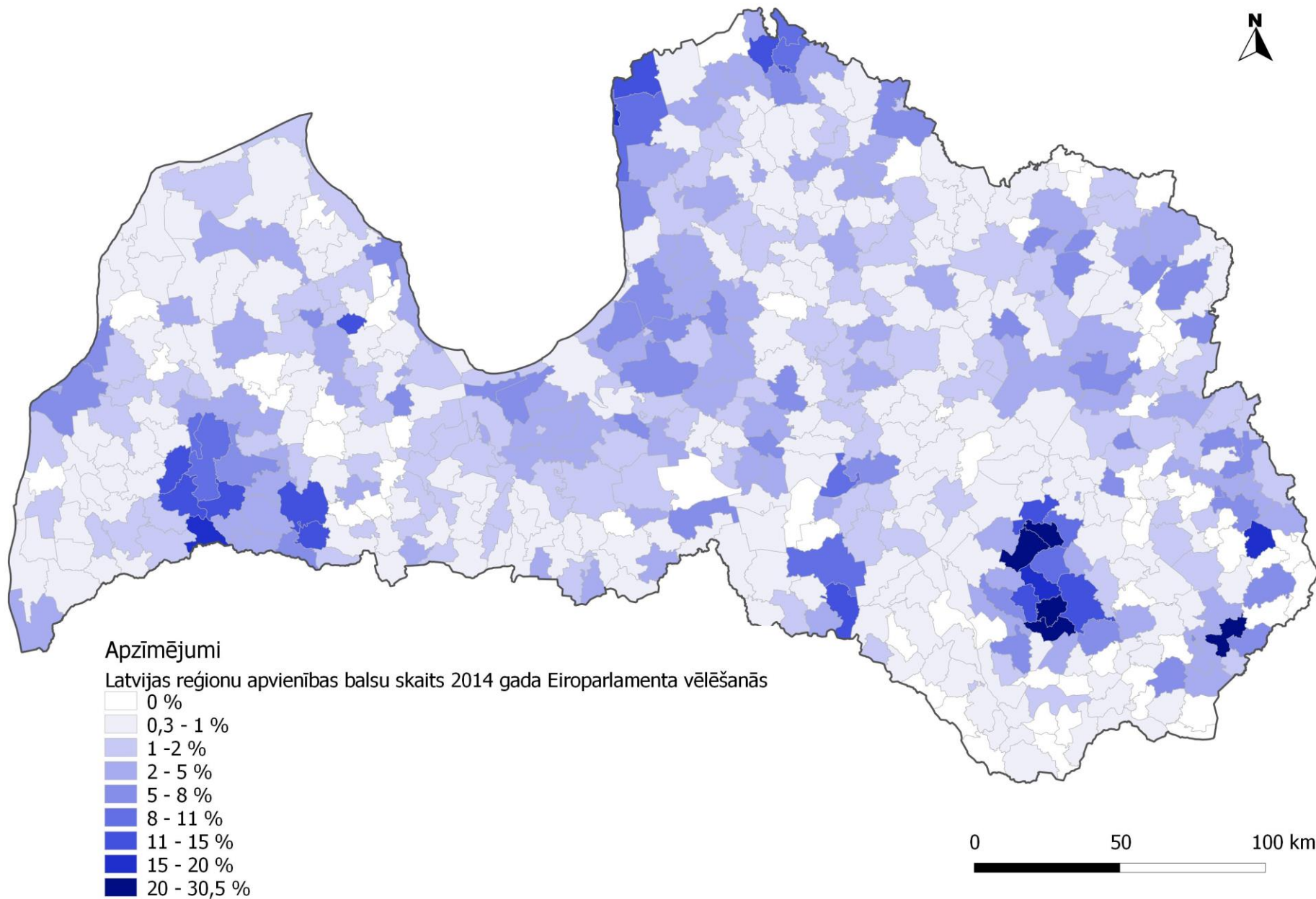
Avots: Jānis Paiders, izmantojot CVK (2014)



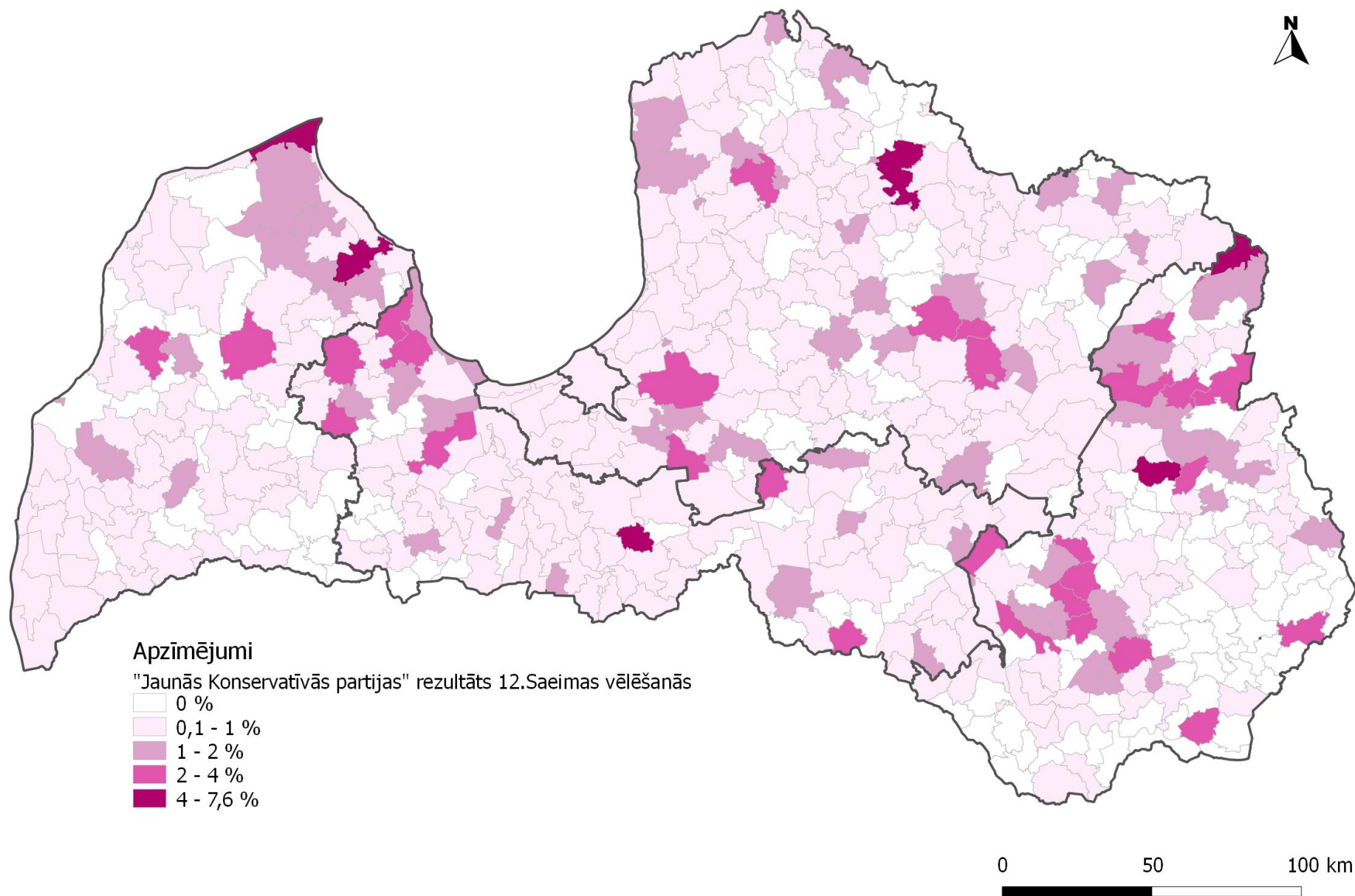
Avots: Jānis Paiders, izmantojot CVK (2014)



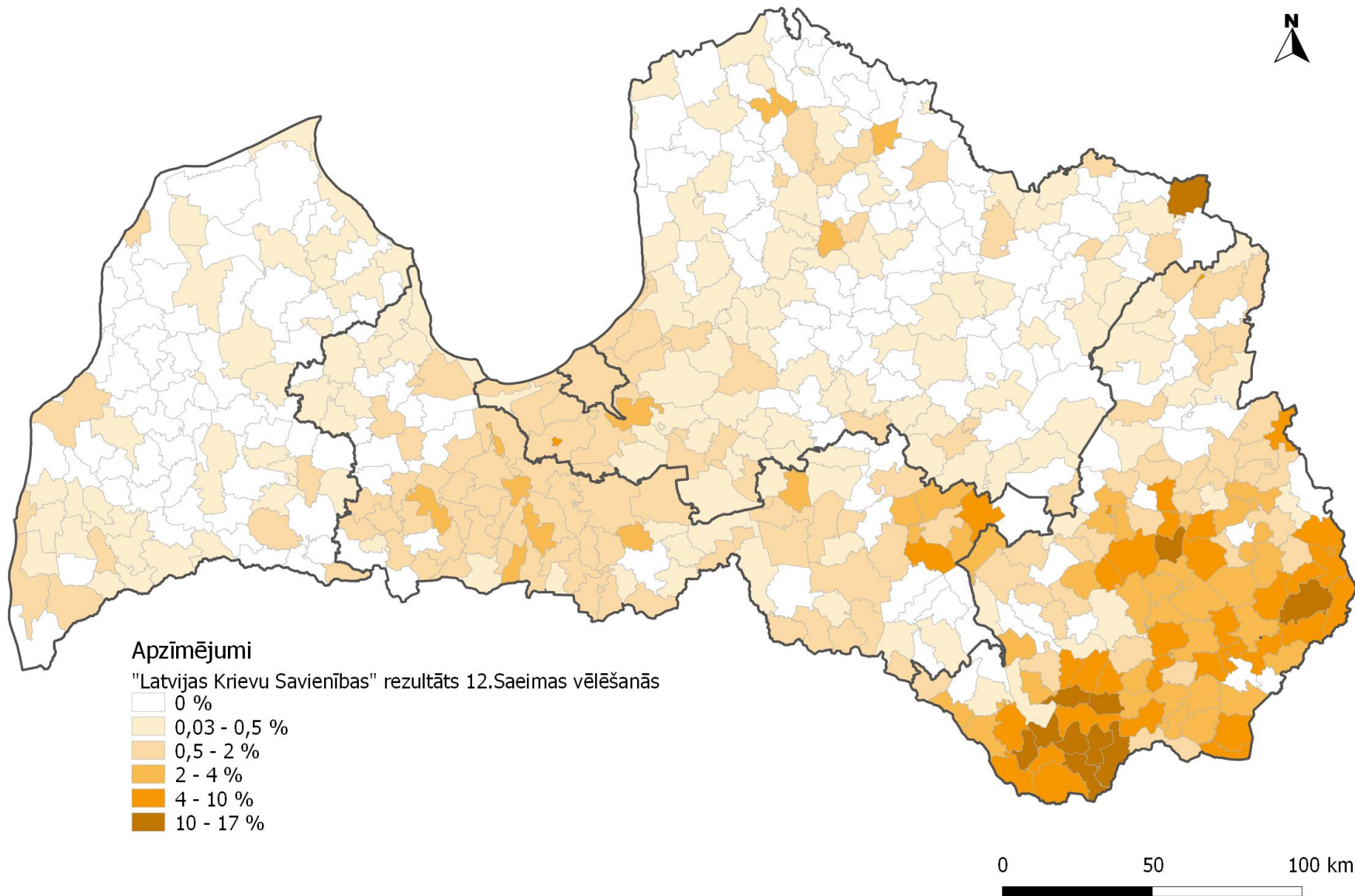
Avots: Jānis Paiders, izmantojot CVK (2014)



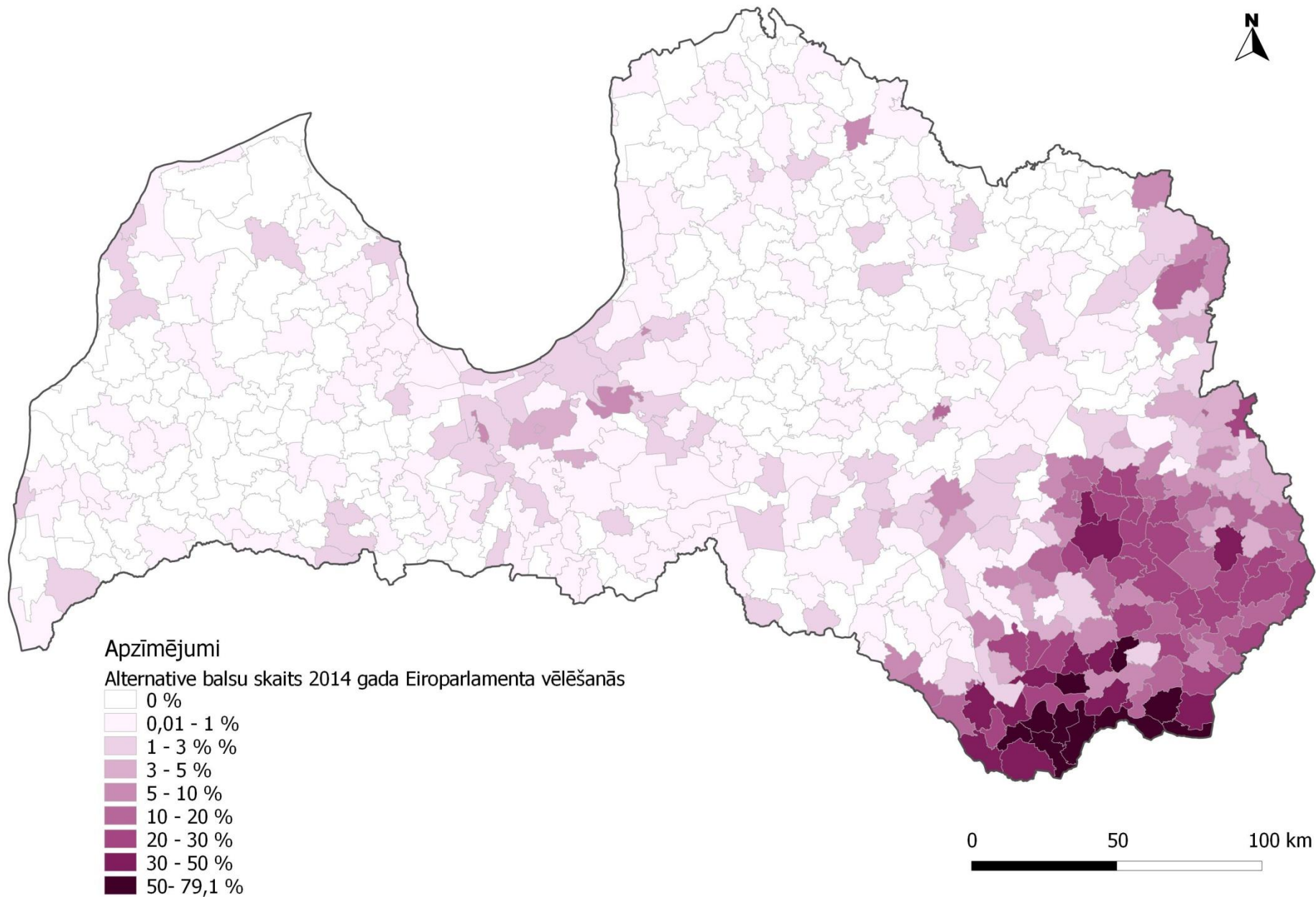
Avots: Jānis Paiders, izmantojot CVK (2014)



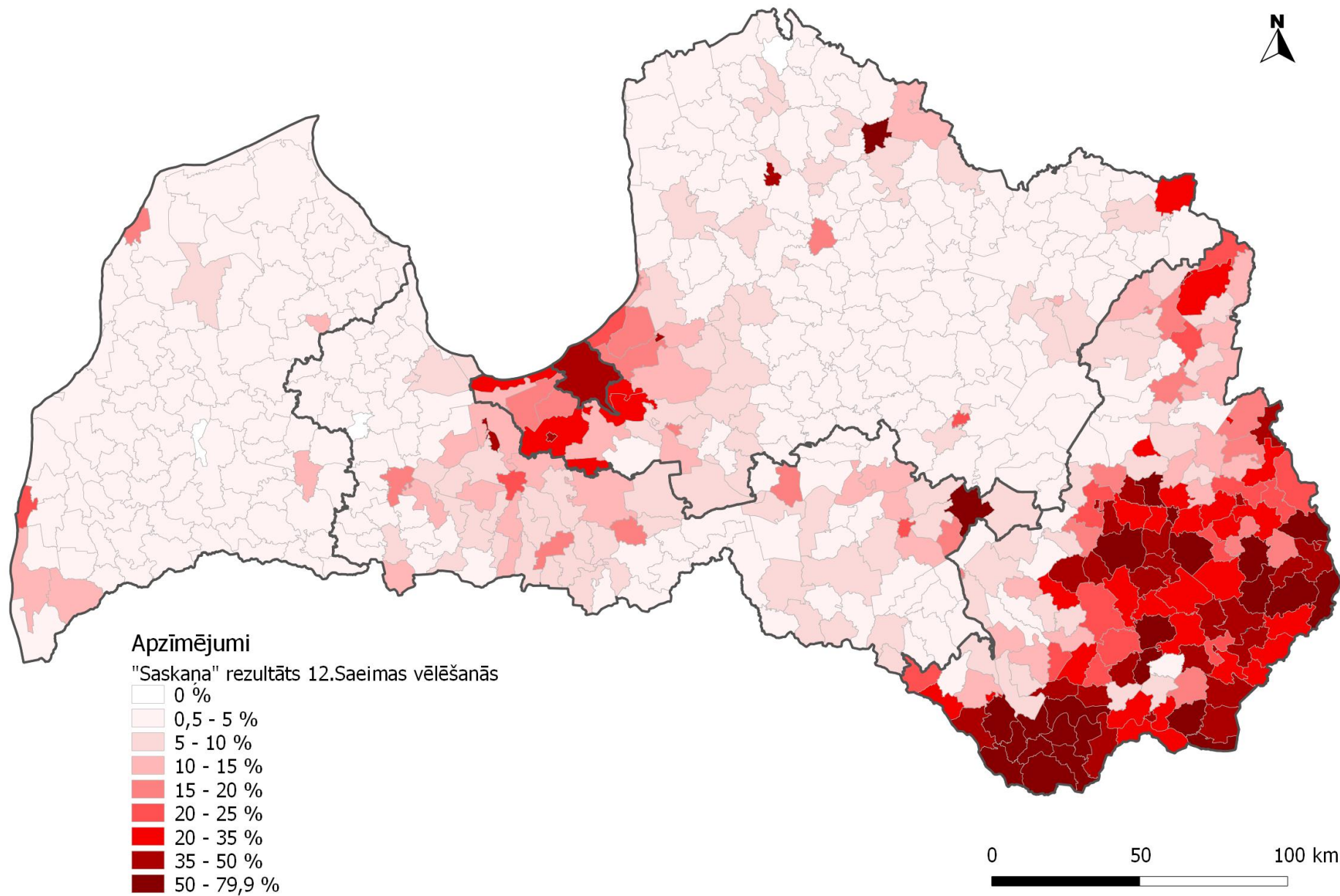
Avots: Jānis Paiders, izmantojot CVK (2014)



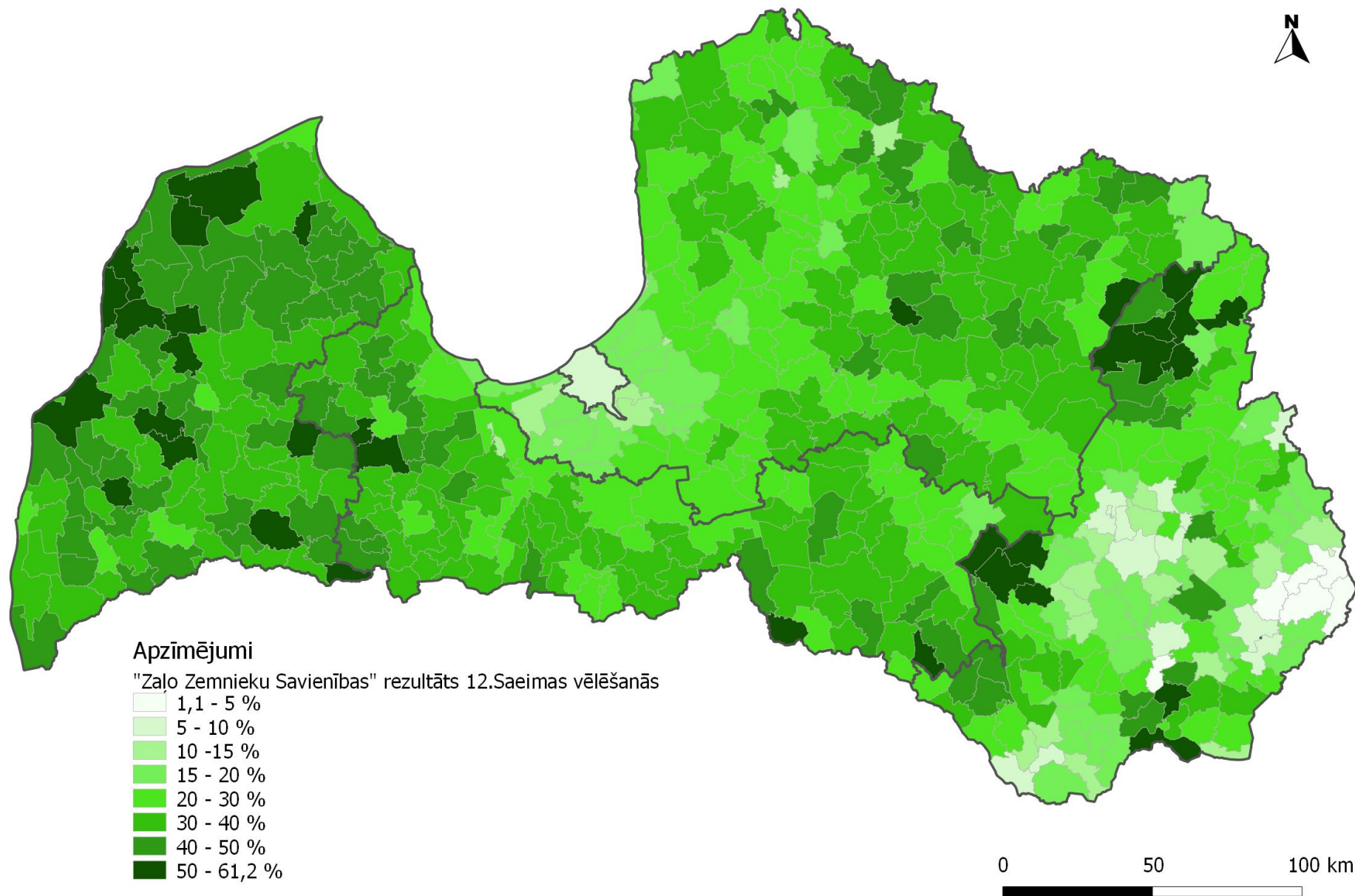
Avots: Jānis Paiders, izmantojot CVK (2014)



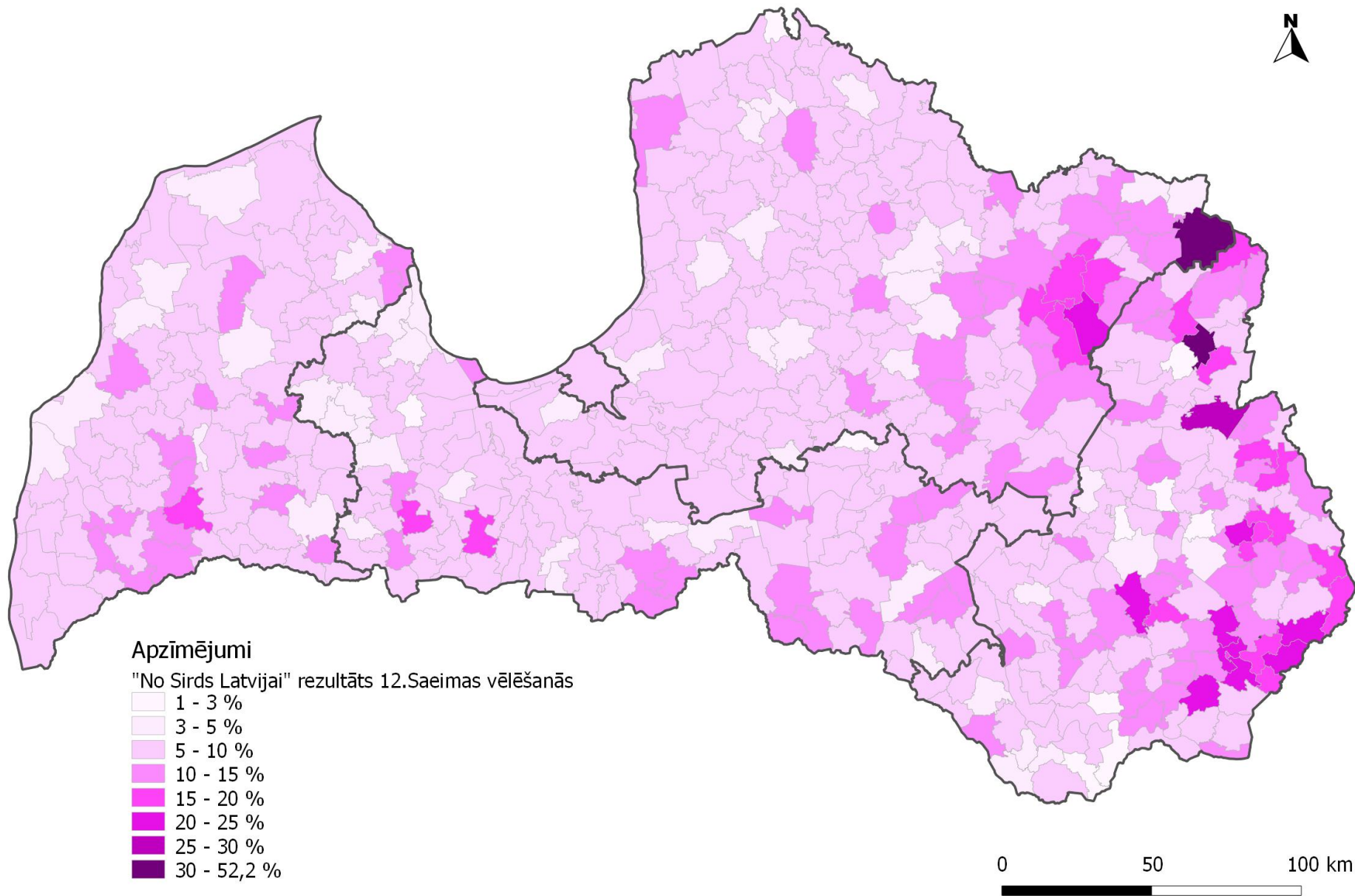
Avots: Jānis Paiders, izmantojot CVK (2014)



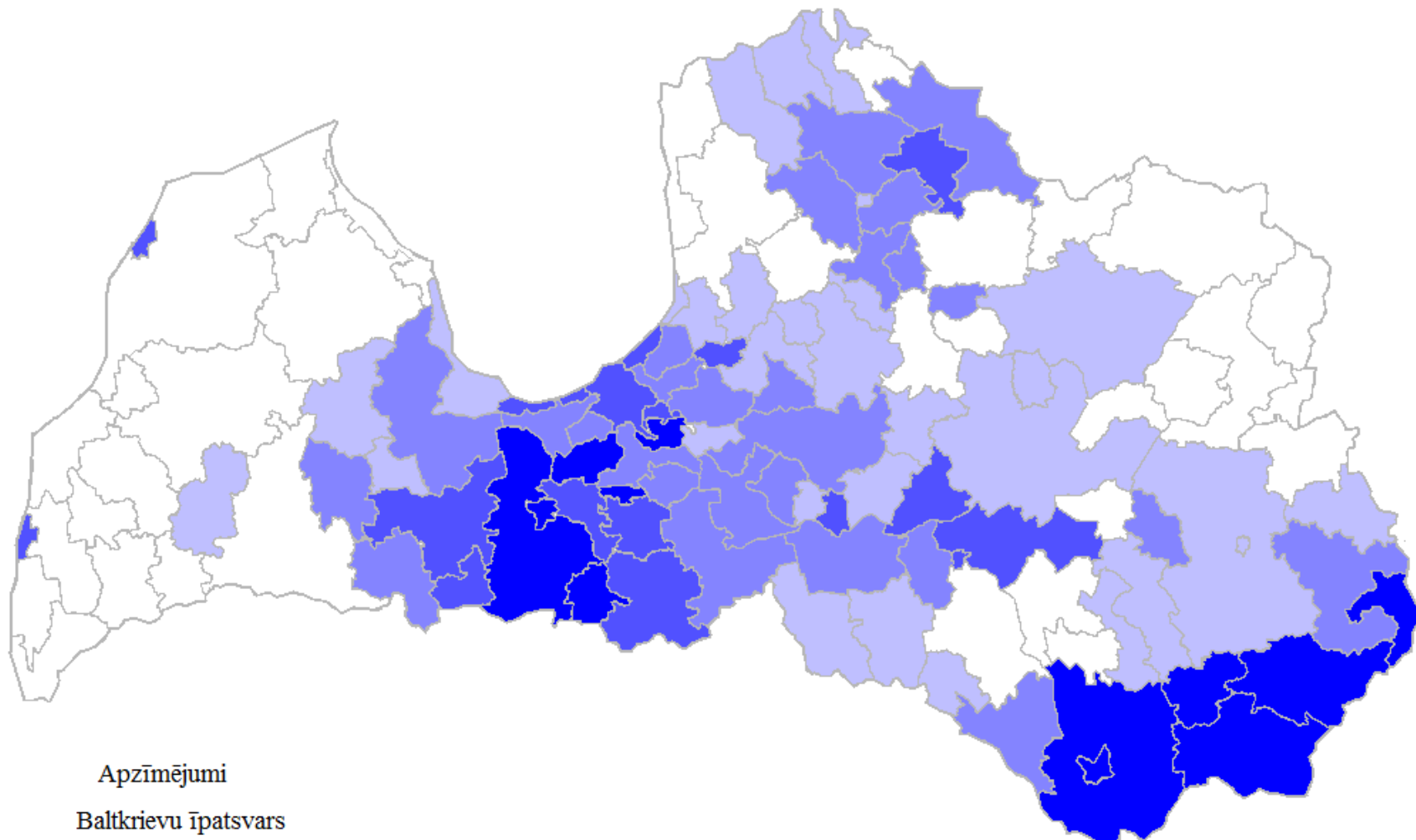
Avots: Jānis Paiders, izmantojot CVK (2014)

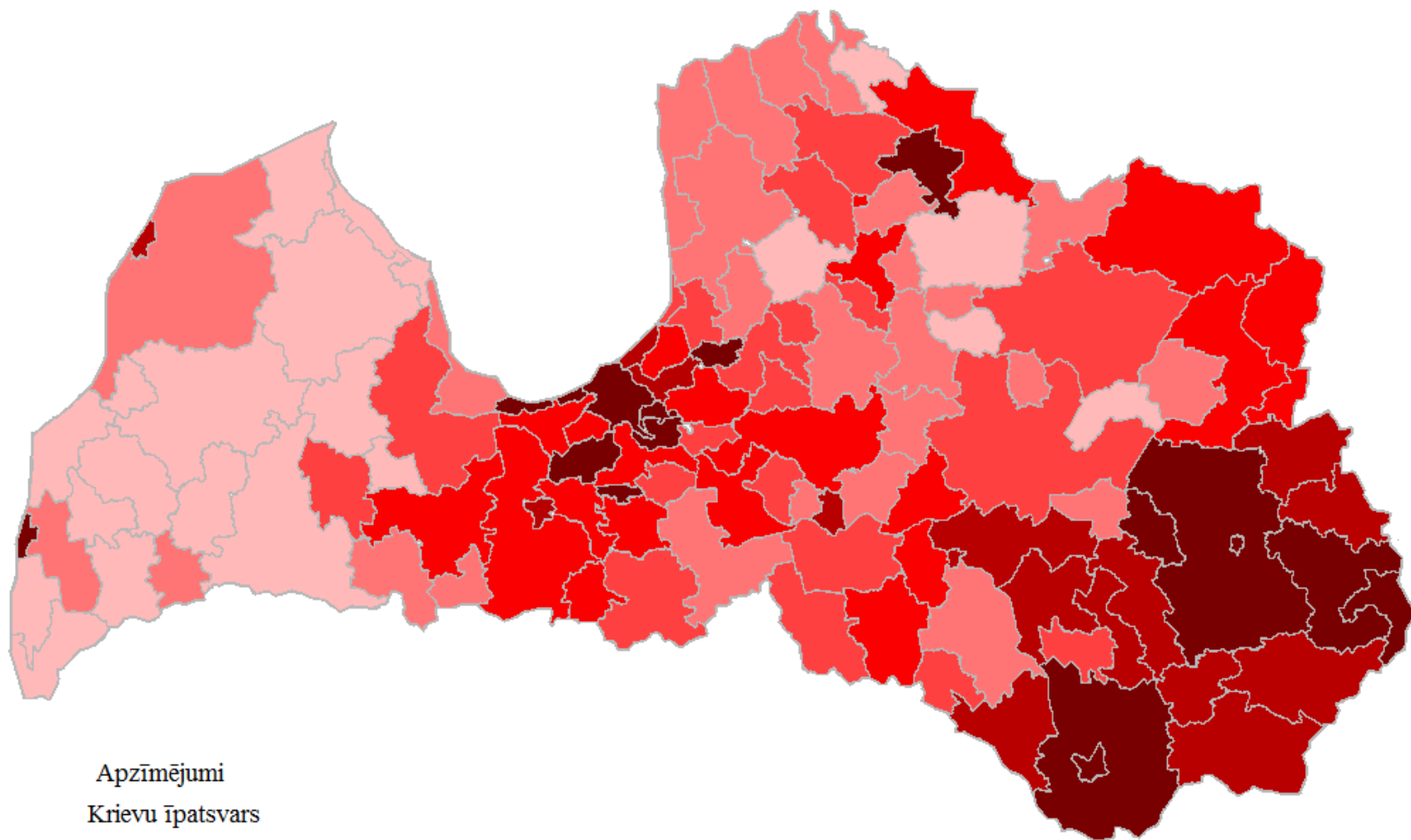


Avots: Jānis Paiders, izmantojot CVK (2014)

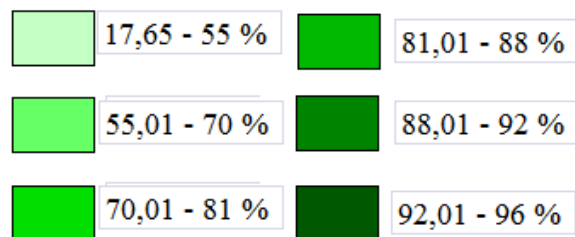
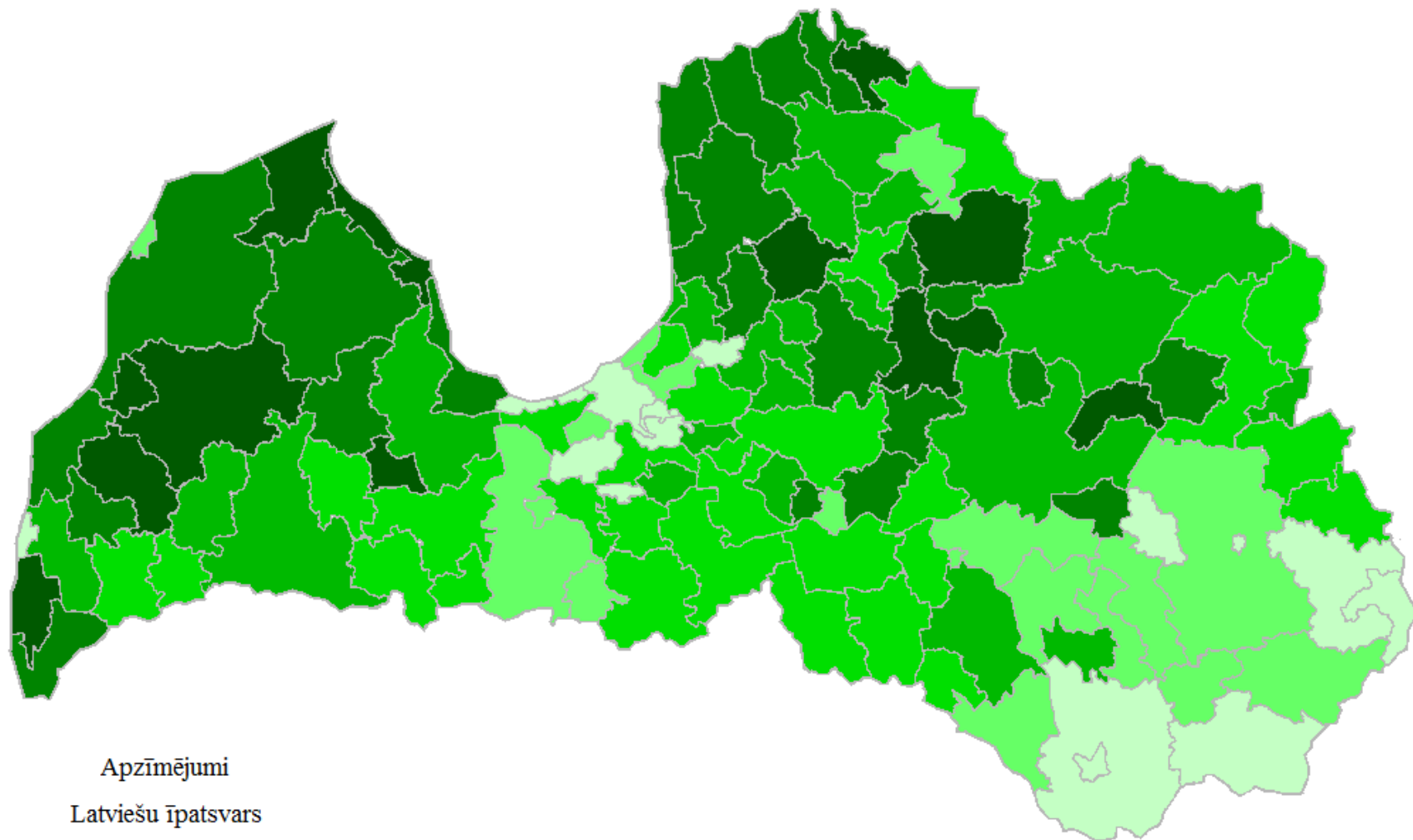


Avots: Jānis Paiders, izmantojot CVK (2014)

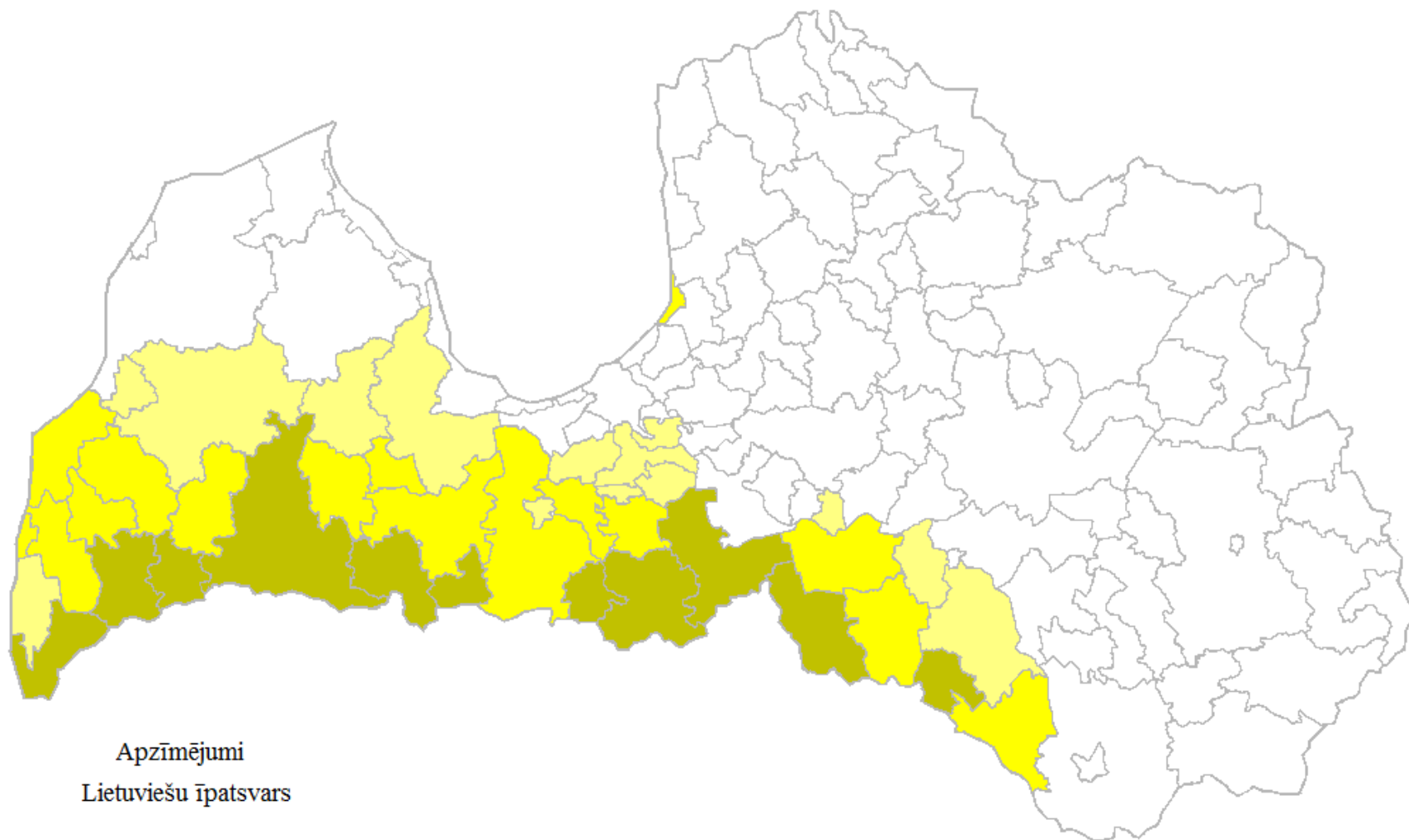


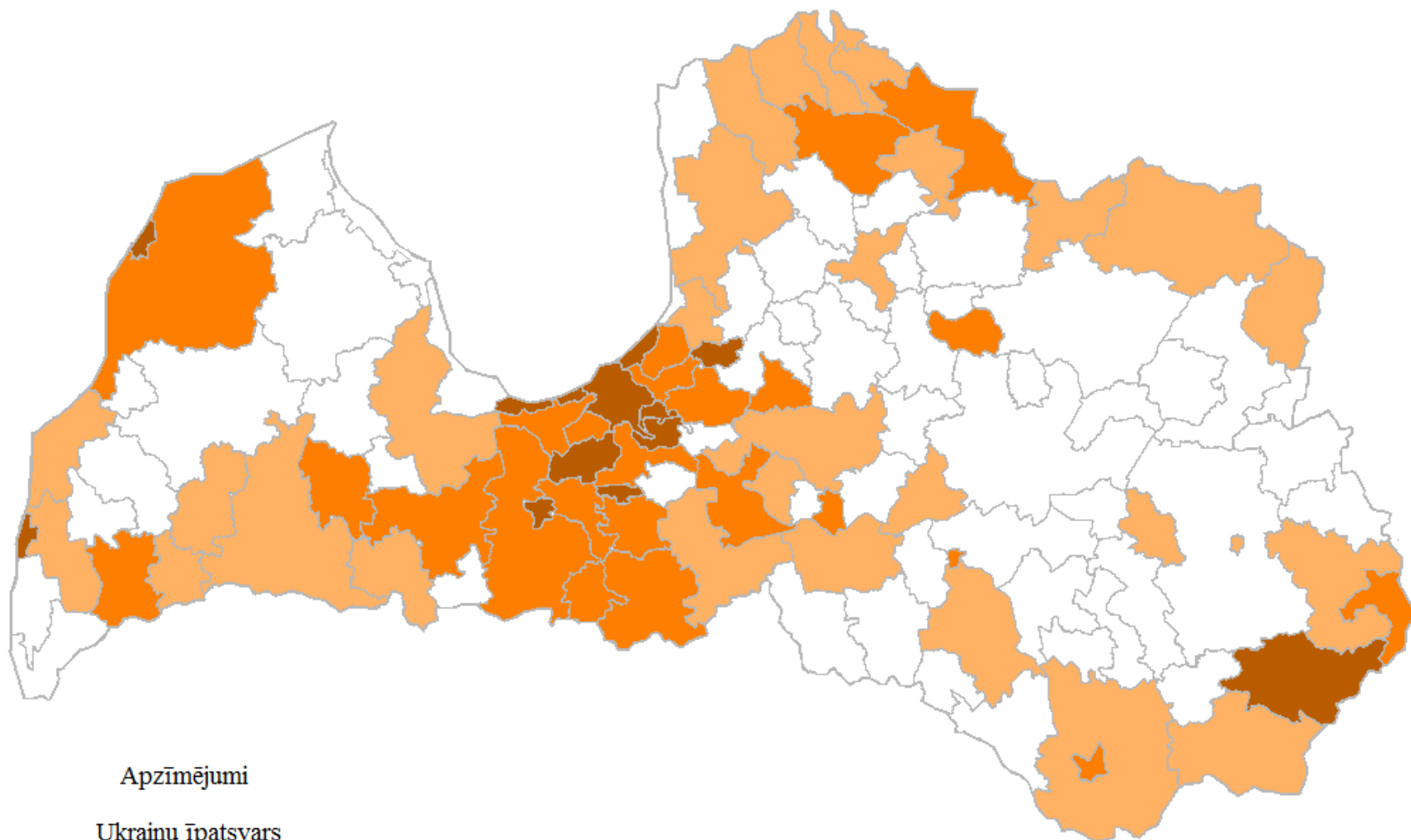


Avots: Jānis Paiders, izmantojot PID datus



Avots: Jānis Paiders, izmantojot PID datus





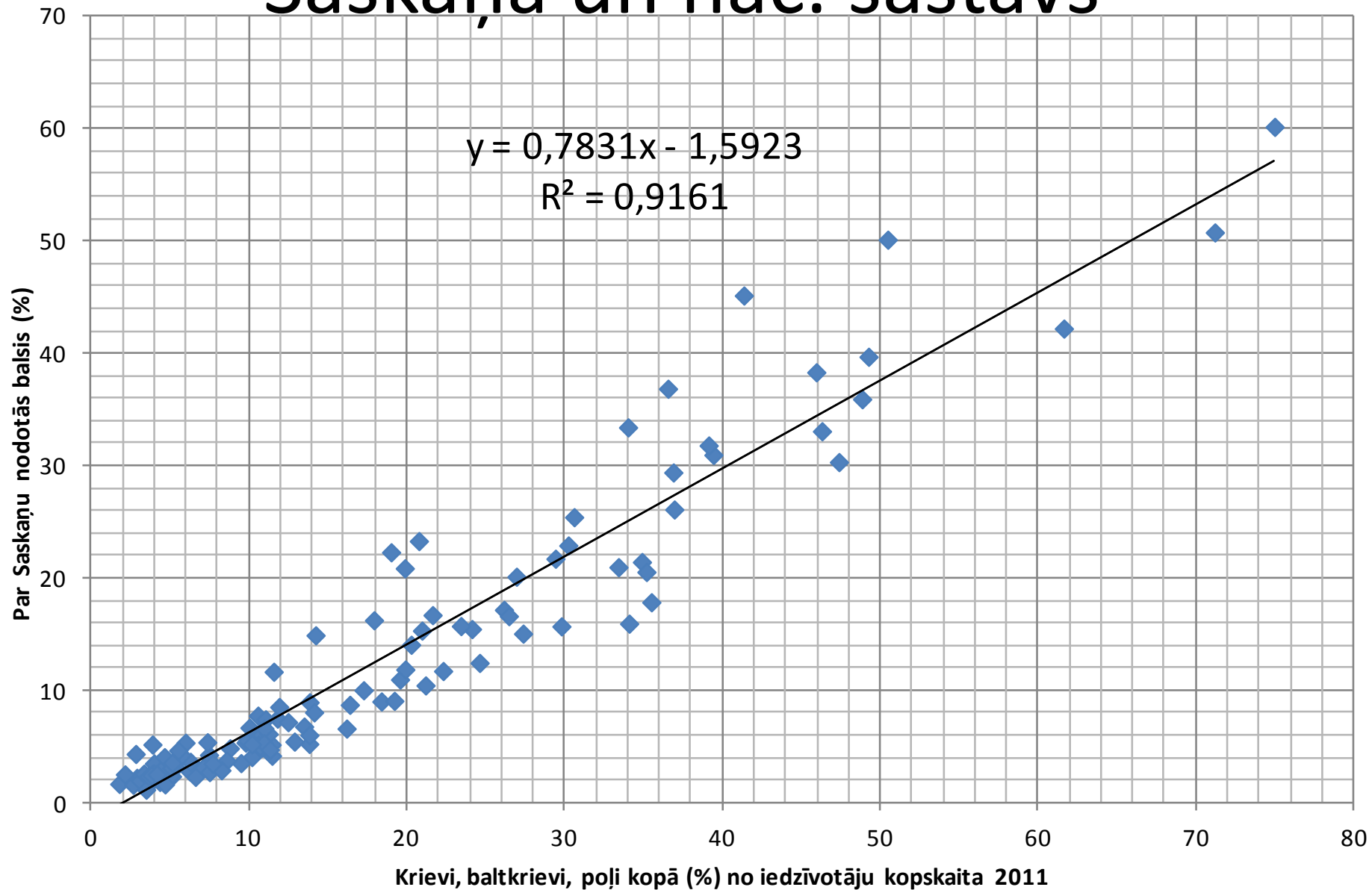
Diemžēl daudzi svarīgi indikatori tiek apkopoti un ir pieejami atbilstoši spēkā esošajam administratīvajam iedalījumam

- Vērtējot novadus un republikas pilsētas, iedzīvotāju skaita atšķirība starp mazāko un lielāko teritoriālo vienību ir mērāma simtos reižu, bet nefinanšu investīciju apjoms var atšķirties jau par tūkstošiem reižu.

Skaitliski lielās nelielo novadu grupas tendences

- Veicot korelācijas koeficienta aprēķinus vai regresijas analīzi ar standarta datorprogrammām, tiek mākslīgi novienādota lielu un nozīmīgu teritoriju ietekme, salīdzinot ar skaitliski lielāku nelielo novadu skaitu. Šādi iegūti statistiskās analīzes rezultāti būtiski samazina pēc iedzīvotāju vai ekonomikas potenciāla lielāko teritoriju ietekmi. Regresijas analīze vairāk izcels skaitliski lielās nelielo novadu grupas tendences, nevis kopējo - visa valsts mēroga tendenci.

Saskaņa un nac. sastāvs



Latvijas pilsoņu nelatviešu
daudzumam pieaugot par vienu
procenta punktu, balsotāju
daudzums par *Saskaņu* un LKS
kopā pieaug par 1 procenta punktu
(ja vēlamies būt precīzi tad par
0,954).

Saskaņas un LKS rezultāti ir funkcionāli saistīti ar nacionālo minoritāšu izvietojumu

- Sociāldemokrātiskās partijas *Saskaņas* rezultāti ir gandrīz funkcionāli saistīti ar tādiem rādītājiem kā krievu, baltkrievu un poļu īpatsvars novadā, vai Latvijas pilsoņu neatviešu īpatsvars novadā. 2014. gada vēlēšanās 92% no balsojuma par *Saskaņu* var izskaidrot ar neatviešu pilsoņu daudzumu attiecīgajā novadā (determinācijas koeficients 0,92).

Veicam pētījumos ar daudzām valstīm

- Vēl lielākas nobīdes ir vērojamas regresijas analīzē, kurā kā datu kopa ir iekļautas visas pasaules valstis. Veicot korelācijas koeficienta aprēķinus vai regresijas analīzi ar standarta datorprogrammām, Ķīnai, Indijai un ASV tiek dots tikpat liels svars kā valstīm ar tikai dažiem miljoniem iedzīvotāju. Šāda analīze parādīs vidējo un mazo valstu tendences, kas var ievērojami atšķirties no kopējo globālo procesu tendencēm.

Diagramma, kurā katrai valstij ir vienāds svars

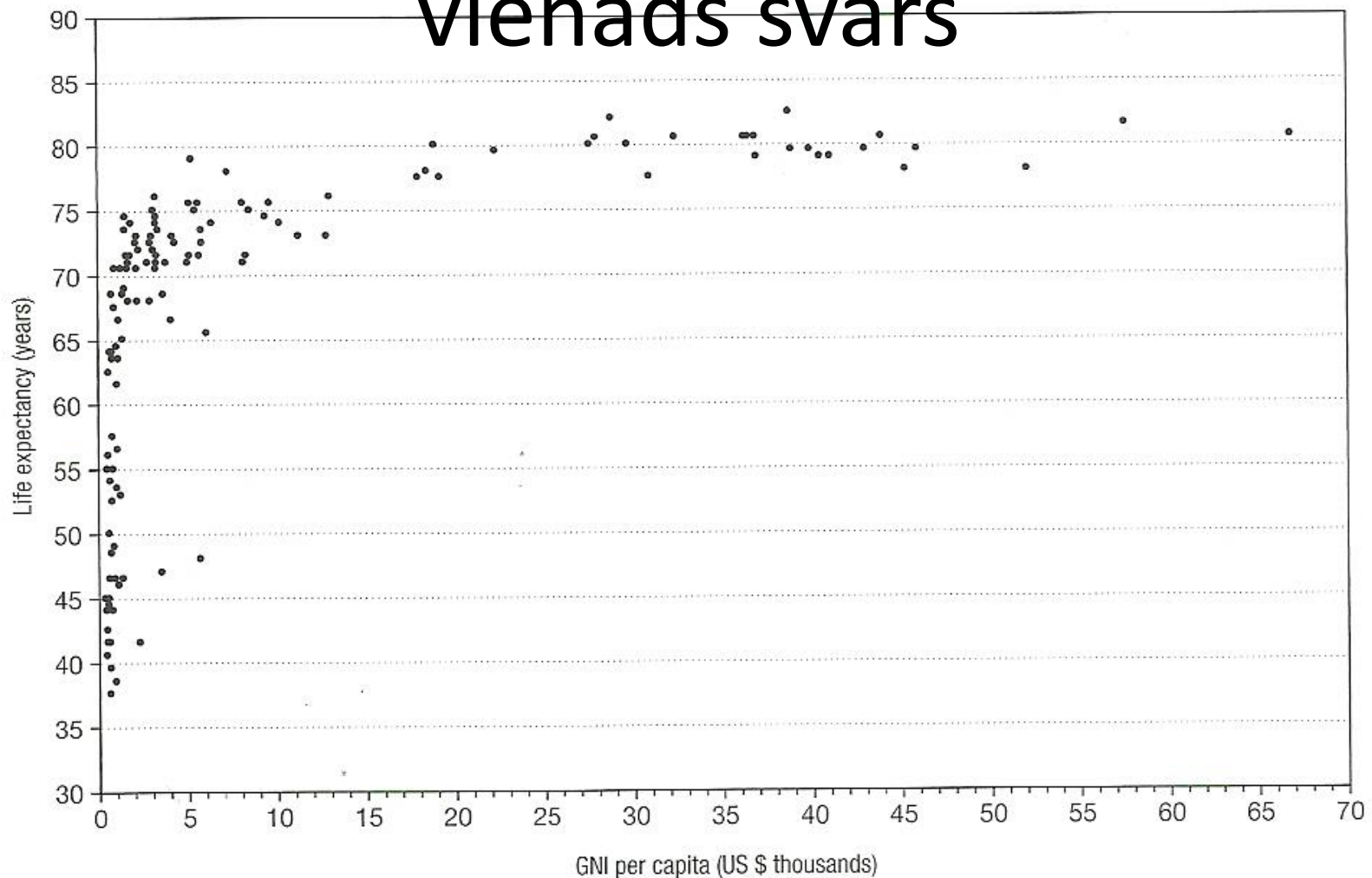


Figure 6.7 GNI per capita and life expectancy

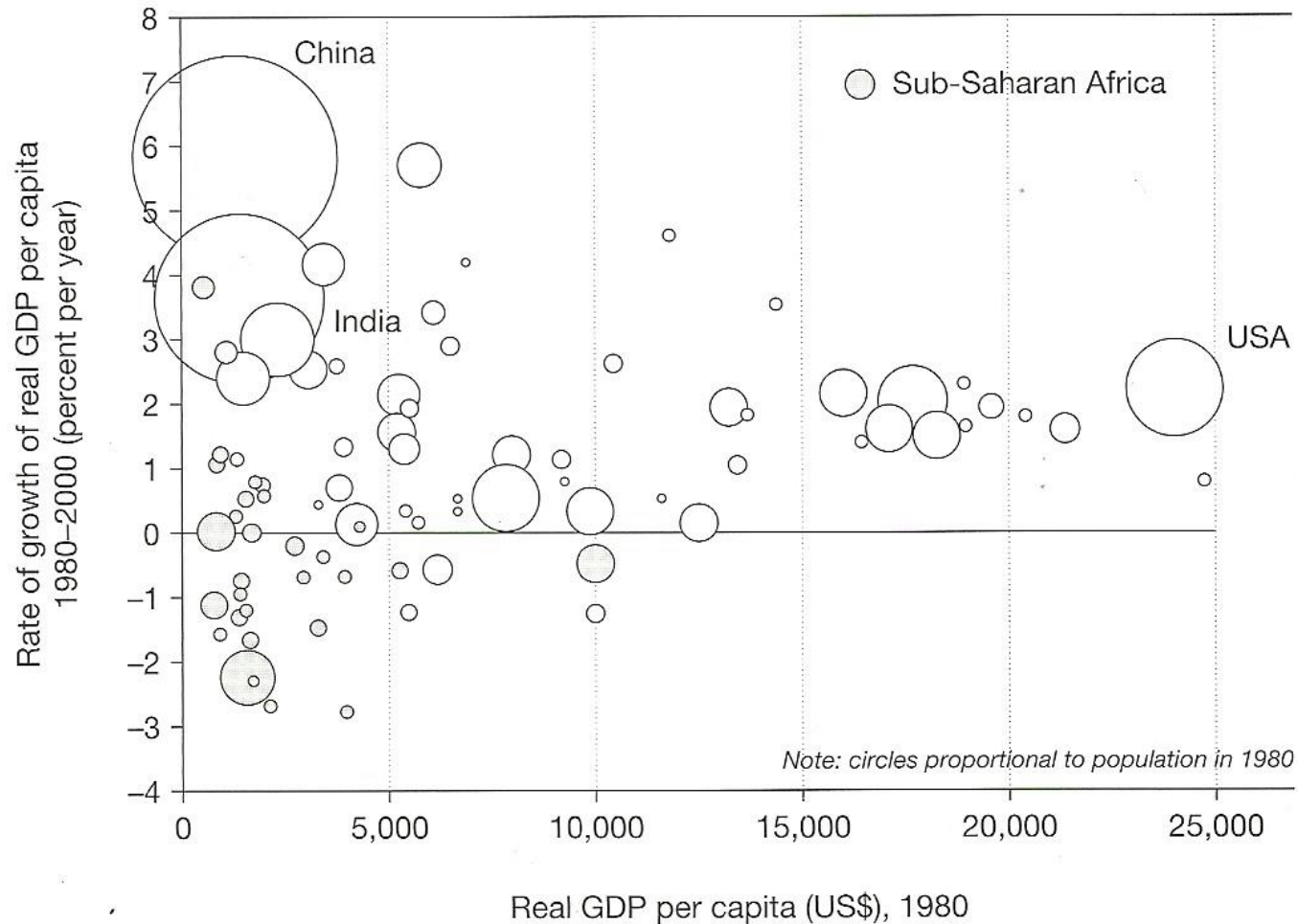
Source: data from World Bank 2009

Lielākām valstīm ir lielāks svars

- Viens no risinājumiem, kā uzlabot matemātisko modeļu precizitāti, ir, rēķinot korelācijas, determinācijas vai regresijas koeficientu, pielietot atšķirīgus svarus lielākām un mazākām teritoriālajām vienībām. Lielākai teritorijai tiktu piešķirts lielāks svars, kas būtu atbilstošs šīs teritorijas nozīmei vai ietekmei.

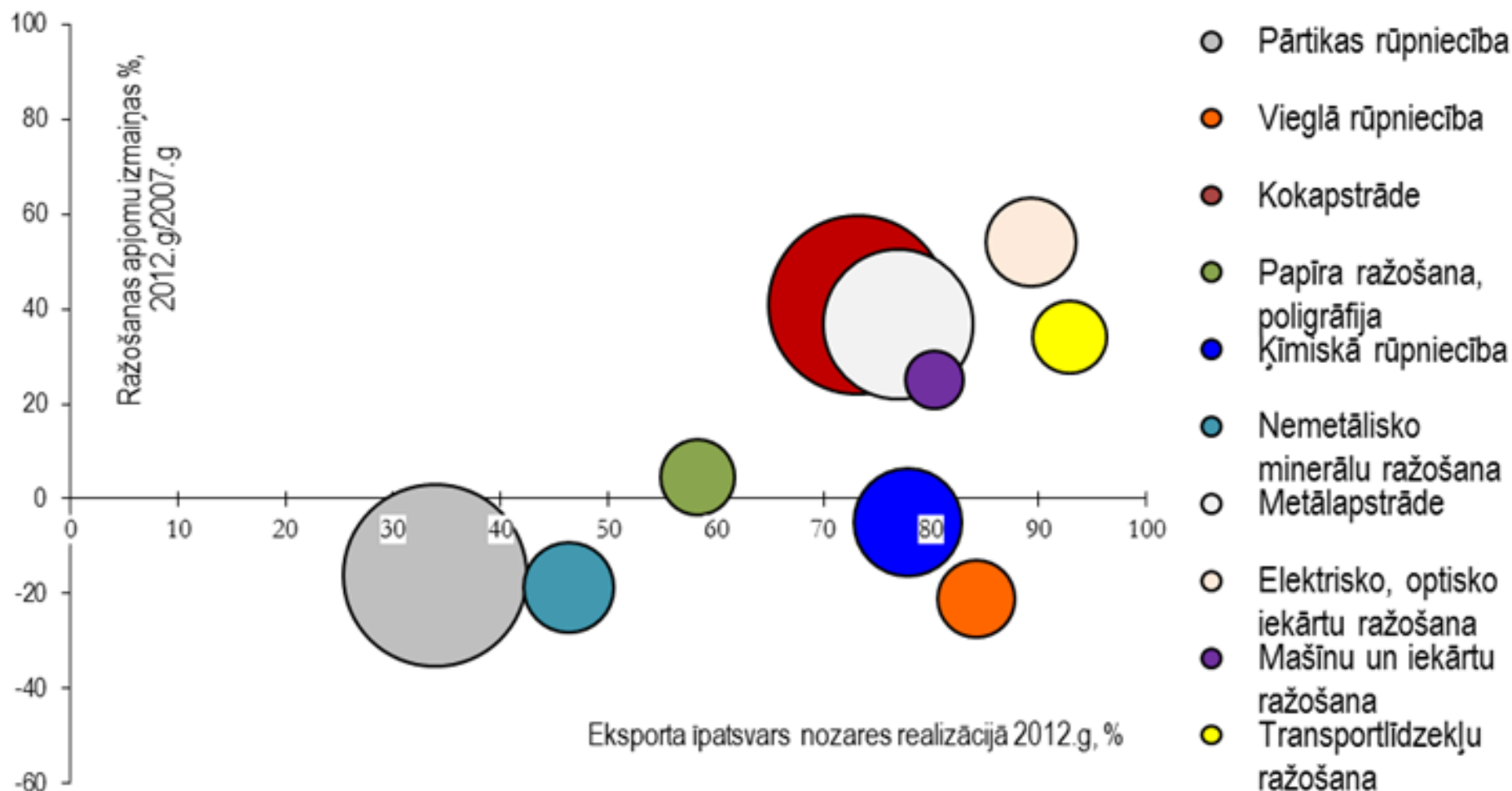
Diagramma, kurā katrai valstij svars ir vienāds ar iedzīvotāju skaitu

114 GEOGRAPHIES OF UNEVEN DEVELOPMENT



Latvijas rūpniecības nozaru potenciāls

Nozares īpatsvars, ražošanas apjomu izmaiņas un eksporta īpatsvars realizācijā²⁹



Nacionālās industriālās politikas pamatnostādnes 2014.-2020.gadam

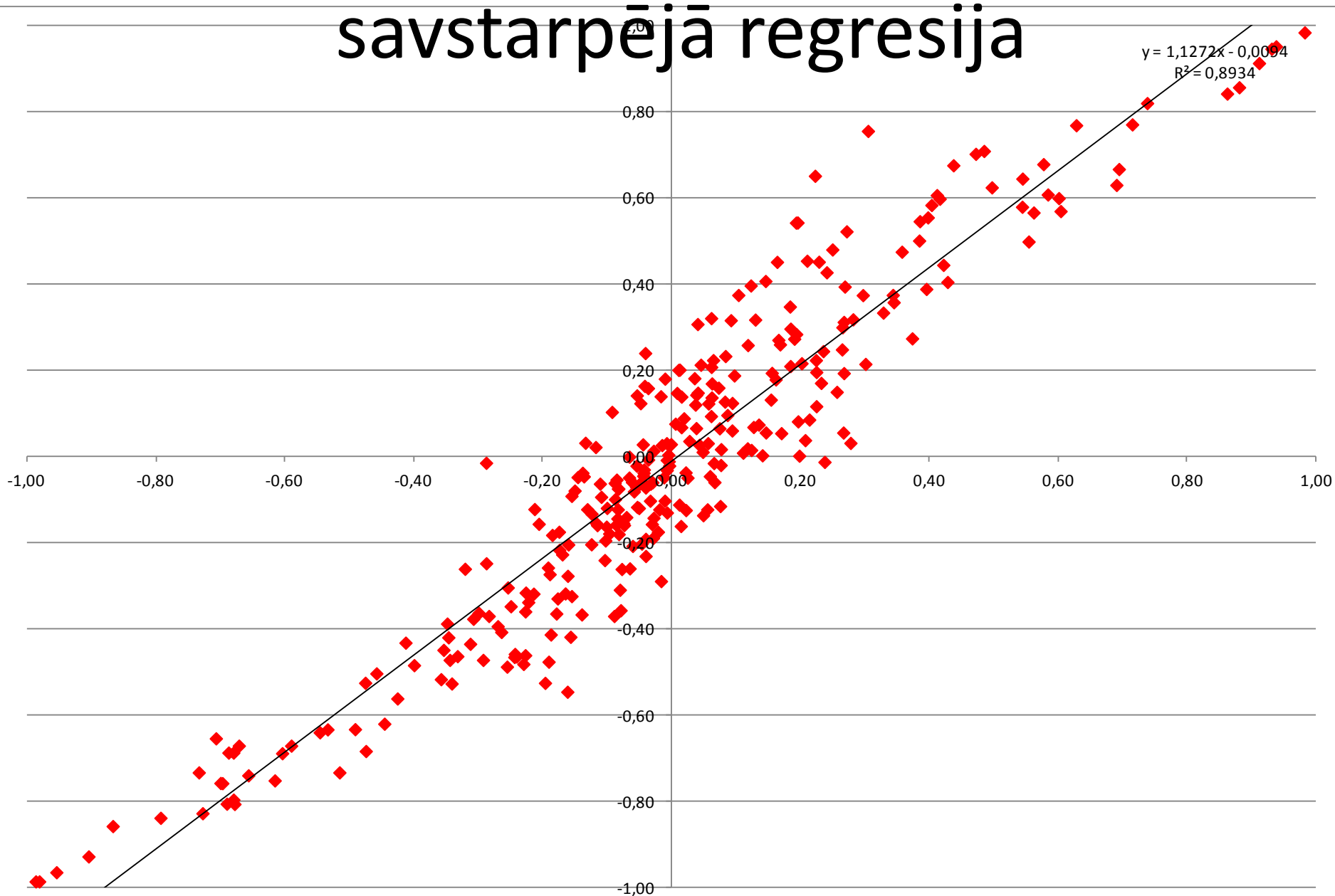
Darba rezultāti

- Autors veica divu korelācijas koeficientu kopu salīdzinājumus, analizējot Latvijas Saeimas vēlēšanu rezultātus. Viena korelācijas koeficientu kopa tika iegūta, izmantojot tradicionālo pieeju, kad katram novadam vai republikas pilsētai ir vienāds svars.
- Otra korelācijas koeficientu kopa tika iegūta, piešķirot katrai administratīvajai vienībai svaru, kas bija proporcionāls vēlēšanu iecirkņu skaitam attiecīgajā administratīvajā vienībā.

Svērtas korelācijas princips - katras teritoriālās vienības rādītāji tika iekļauti tik reizes, cik vēlēšanu iecirkņu bija attiecīgajā administratīvajā vienībā.

Šāda pieeja ļauj aprēķināt korelācijas, determinācijas un regresijas koeficientus, dodot lielāku svaru administratīvi teritoriālajam vienībām ar lielāku vēlēšanu skaitu

Divu korelācijas koeficientu matricu savstarpējā regresija



Uzlabošanas iegūtā regresijas modeļa precizitāte

- Pielietojot šādu pieeju, uzlabošanas iegūtā regresijas modeļa precizitāte, taču atšķirīgu svaru pielietošanas gadījumā regresijas formālais novērojumu skaits pieaug, jo vieni un tie paši novērojumi teritorijai ar lielāku svaru tiek iekļauti vairākas reizes, tāpēc **regresijas koeficientu ticamības intervāla būtiskuma radītāji tiek mākslīgi palielināti.**