

# RÉGIÓDETEKTÁLÁS HOSSZÚTÁVÚ NÉPESSÉGI ADATOK ALAPJÁN

---

**MÁTÉ GABRIELL**

BABEŞ-BOLYAI TUDOMÁNYEGYETEM  
FIZIKA KAR, FIZIKA-INFORMATIKA SZAK,  
4. ÉVFOLYAM

TÉMAVEZETŐK

**DR. BENEDEK JÓZSEF**

BABEŞ-BOLYAI TUDOMÁNYEGYETEM  
FÖLDRAJZ KAR, TÁRSADALOMFÖLDRAJZ TANSZÉK

**DR. NÉDA ZOLTÁN**

BABEŞ-BOLYAI TUDOMÁNYEGYETEM  
FIZIKA KAR, ELMÉLETI ÉS SZÁMÍTÓGÉPES FIZIKA TANSZÉK

XI. ERDÉLYI TUDOMÁNYOS DIÁKKÖRI KONFERENCIA

KOLOZSVÁR, 2008. MÁJUS 23 – 24

## Kivonat

Napjainkban egyre fontosabb feladatnak számít, úgy szociális, mint gazdasági szempontból, régiók objektív módon való definiálása. Ez általában félempirikus módszerekkel és sok paraméter figyelembevételével (inkább ráérzéssel) történik. Alkalmazva a statisztikus fizika sok más területén is jól bevált modelljeit és módszereit, két eljárást is sikerült kidolgoznunk, melyek segítségével különböző skálákon definiálhatunk-detektálhatunk régió-szerű klasztereket. Módszerünk hosszútávú népszámlálási adatokra illetve az ezen belül megfigyelhető települések közti korrelációk detektálására épül. A jelen dolgozatban ezen módszerek alapelve és egyszerű alkalmazása kerül bemutatásra.

## Tartalom

1.	Bevezető.....	3
2.	Adatok.....	5
3.	A módszer .....	5
3.1.	A determinisztikus módszer .....	7
3.2.	A sztochasztikus módszer .....	9
4.	Tárgyalás és Következtetések .....	14
5.	Hivatkozások.....	16
6.	Függelék.....	18

# 1. Bevezető

A XX. század második felétől egyre több tudományágból, egyre több tudós fordítja figyelmét a régiók kutatása felé. Ugyanakkor a politikusokban is egyre jobban tudatosul ezek jelentősége gazdasági, szociális, és kulturális szempontokból egyaránt, ismételten motiválván ezzel a kutatókat [1,3]. Gazdasági szempontból a legfontosabbnak a funkcionális régiók tekinthetők. Mi magunk is, mindennapi teendőink során, (ezeknek megfelelően) régiókat definiálunk magunk körül, ezáltal könnyebbé téve életünket. Például, ha csak egy kenyérre van szükségünk, nyilván a „régió”, ahol a kenyeret keressük, elég kicsi lesz, ha viszont egy televízió készüléket szeretnénk vásárolni, hajlamosak vagyunk ezért nagyobb távolságot megtenni, több üzletet felkeresni, annak érdekében, hogy minél olcsóbban jussunk, minél jobb minőséghez. A fenti példa esetén is két jól meghatározott objektum köré építettük régióinkat. Az egyik maga a lakhelyünk – a központ, ahova a vásárolt termék kerül, a másik pedig az üzlet, ahonnan az adott terméket megvásároltuk – mondjuk úgy, ennek származási helye. A funkcionális régiók tulajdonképpen két hasonló fogalom alapján értelmezhetők: egy (vagy több) központ (vagy csomópont), és az ehhez valamilyen szempontból hozzá tartozó, körötte található terület [2]. Attól függően, hogy milyen célból definiáltuk a funkcionális régiót, a csomópont lehet más és más. Leggyakrabban ezen régiók megállapítása gazdasági megfontolásokon alapszik, ugyanis az egy régióba tartozó területek általában gazdaságilag együtt működnek, munkaerő-piacuk közös, ugyanakkor a régió határain belül való kereskedés jövedelmezőbb [2,3]. Az előbb felsoroltakból is felmérhetjük, miért is fontos, hogy minél helyesebben és pontosabban meghatározzuk ezeket a régiókat. A jelenleg rendelkezésünkre álló módszerek ezen régiók feltérképezésére nem adnak egy pontos, objektív megoldást; intuitívak, a kutatók sokszor megérzéseikre kell hagyatkozzanak. Néhány modell a munkaerőre, és a munka-kínálatokra építkezve, az ingázók száma alapján állapítja meg a régiókat [2,3]. Ezen modellek hiányossága azonban az, hogy

nem veszik figyelembe, sem a különböző képezéseket (szükségletben és kínálatban egyaránt), sem a jövedelemben való különbözőségeket két centrum között [2].

Jó volna találni egy olyan metódust tehát, ami kiküszöböli (kikerüli) ezen hibákat. Az hogy milyen irányból közelítettük meg a megoldást, látszólag talán kissé távol áll azon szakterületektől (földrajz, közgazdaságtan, demográfia, szociológia, stb.) amelyek hagyományosan a régiók kutatásával foglalkoznak. Amit használunk, az fizika. Nyilván sokan felteszik magukban a kérdést, hogy hogyan téved egy fizikus a fizikától (ismételten látszólag) ennyire távol álló területre. Tudni kell azonban, hogy a fizikát manapság nagy sikerrel alkalmazzák olyan tudományokban, mint közgazdaságtan, szociológia, nem beszélve a biológiáról, ökológiáról... Hogy mi is ennek az oka? Talán az, hogy a fizika az egyik (ha nem AZ) alaptudomány, amelyre aztán a többi tudomány építkezik. Mivel ennyire az alapoknál van, a fizika az univerzalitások megállapítására törekszik, azaz megtalálni mindenben azt, ami közös. Ezen közös dolgok alapján modelleket alkotni, és akkor ezek a modellek érvényesek lesznek mindenre. A helyzet persze ennél azért bonyolultabb, mégis, amint látni fogjuk az általunk kidolgozott két eljárásban is, hogy ezek a modellek és módszerek, analógiák alapján, nagyon sikeresen alkalmazhatók. Előnyünkre szolgál, hogy eljárásaink teljesen más alapokon fekszenek, mint a közgazdaságtanban, vagy földrajzban kidolgozott modellek, kiküszöbölve ezzel együtt ezeknek hibáit. Emellett a módszerek alapjainál fekvő gondolatok, modellek, és módszerek már számtalanszor voltak sikeresen alkalmazva, és ezen alkalmazások esetén az elméleti és kísérleti eredmények jó egyezést mutattak, így bízhatunk helyességükben.

Említettem tehát, hogy két módszert sikerült kidolgoznunk. Jelen dolgozatban ezen két módszer felépítése és működése, valamint az ezek által szolgáltatott eredmények kerülnek bemutatásra. Nyilván régiók „detektálásakor” lényegében valamilyen adat feldolgozásáról beszélünk. Mi részletes, időben hosszan visszanyúló népszámlálási adatok használtunk ilyen célból, de érdemes megemlíteni, hogy használhatunk bármilyen típusú adatokat, azzal a feltétellel, hogy sok „időpillanatban” legyen mérésünk, a mérések közötti időintervallumoknak, az adatoknak megfelelően, elegendően nagyoknak és lehetőleg azonos nagyságúaknak kell lenni, a legfontosabb feltétel pedig, hogy két adott település esetén, ezen adatok korrelációja jellemezze a két település közötti kapcsolat erősségét.

## 2. Adatok

Amint említettem, a módszerek működésének szemléltetésére időben hosszan visszanyúló népszámlálási adatokat használtunk. Ezen adatok mellett, minden esetben, szükségesek még a települések földrajzi koordinátái is. A dolgozatban, Erdély esetén, **Varga E. Árpád** budapesti könyvtáros által összegyűjtött, és *Erdély etnikai és felekezeti statisztikájában* [4] közölt adatok alapján dolgoztunk. Itt összesen 6044 erdélyi településre, 24 „időpillanatra” találunk népszámlálási adatokat. Sajnos sok ilyen időpillanatra, az adatok nem népszámlálásból származik, hanem helységnévtárakból és egyéb egyéni munkákból ezért, sok településre, ezek az adatok hiányosak. Mindemellett, népszámlálások esetén is a nagyon kicsi településekre vonatkozó adatok is hiányosak. Első lépésben ezért kivettük az adatbázisból azokat az éveket, melyekre az adatok nagymértékben hiányoztak, ez után pedig eltekintettünk azoktól a településektől, amelyre akár egy hiányos év is akadt (a megmaradt évekből). A települések koordinátáit egy web-fórumon [5] találtuk. Miután metszettük a koordinátákat tartalmazó adatbázist a népszámlálási adatokat tartalmazóval, 3577 településre, időben visszamenőleg 11 évre maradt adat (1880, 1890, 1900, 1910, 1920, 1930, 1941, 1956, 1966, 1977, 1992). A magyarországi adatokat **Nemes-Nagy József** professzor úr (ELTE, Budapest) szolgáltatta. Ez az adatbázis már tartalmazta a települések koordinátáit is, emellett nem voltak olyan évek, amelyek esetén nagyszámú településnek hiányzott volna adata, ezért egyetlen dolgunk az volt, hogy kivegyük azokat a településeket, amelyek esetében valamely évre nem volt adat. Így a kezdeti 3147 településből 3102 települést tudtunk megtartani. Itt is 11 időpillanatra van népszámlálási adat (1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970).

## 3. A módszer

Ahhoz, hogy nekiláthassunk az adatok feldolgozásának, még némi simításra szükség van, nem szabad ugyanis megfélekednünk arról, hogy az a nagyjából 100 év

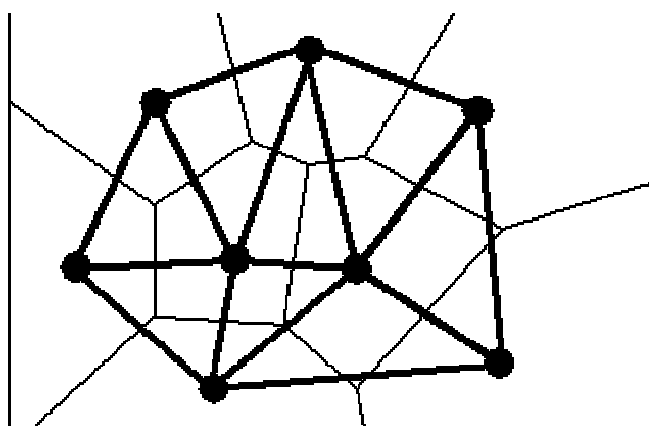
alatt, amire adatunk van, igencsak nagy volt a természetes szaporulat. Mi nyilván nem ezt akarjuk vizsgálni, ezért első lépésben összes település populációját relatív értéként tekintettük, Erdély illetve Magyarország összlakosságához viszonyítva őket (lakosságának számát leosztottuk az adott terület, - Erdély illetve Magyarország - adott évre jellemző összlakosságával.). Ezáltal kiküszöböltük a természetes és triviális növekvő trendet. Az így kapott relatív nagyságot ezentúl egyszerűen nagyságnak ( $r$ ) nevezzük, és minden egyes települést ezzel jellemezzük.

Rendelkezésünkre állnak tehát azon adatok, amivel a településeinket tudjuk jellemezni. Ahhoz, hogy régiókat tudjunk detektálni, szükséges először, hogy meghatározzuk a települések közt fennálló szomszédsági relációkat (vagyis kinek kik a szomszédjai), a régiók ugyanis ezen szomszédok által válnak összekapcsolttá. Geometriailag tekintve, a települések egy síkbeli ponthalmazt alkotnak. Minden egyes pontnak pontosan ismerjük a koordinátáit és célunk ezen ponthalmaz elemeit összekötni, úgy, hogy szomszédsági relációkat értelmezzünk. Ennek érdekében a ponthalmazhoz rendelt Voronoi felosztást használjuk (1. ábra). Értelmezés szerint, a tér egy  $S$  ponthalmazának Voronoi diagramját [6,7] megkapjuk, ha  $S$  minden eleme esetén meghatározzuk a tér azon  $V(p)$ ,  $p \in S$  tartományait, úgy, hogy

$$d[p, k] \leq d[q, l], \forall k, l \in V(p)$$

$$p, q \in S, p \neq q,$$

ahol  $d[A, B]$  az  $A$  és  $B$  pontok Euklideszi távolsága. Azaz  $V(p)$  a tér azon pontjainak halmaza, melyek legközelebbre a  $p$  ponthoz találhatók  $S$ -ből (esetleg létezik  $S$ -nek más olyan pontja, melynek távolsága  $V(p)$  valamely pontjától megegyezik a  $p$  pontnak az ugyanettől a ponttól mért távolságával). Két települést szomszédosnak fogunk tekinteni, ha az általuk meghatározott Voronoi cellák határosak.



1. ábra, példa a sík Voronoi felosztására; a vastag sötét vonallal összekötött pontok, a szomszédos Voronoi celláknak megfelelő pontok

Miután meghatároztuk mely települések szomszédosak, lássuk most, hogyan tudnánk meghatározni ezen szomszédsági kapcsolatok erősségét. Fontos, hogy egy jó mérőszámot találjunk ennek a kapcsolat-erősségnek, hiszen valami módon, ez alapján tudjuk majd eldönteni, hogy két település egy régióba tartozik vagy sem. Tekintve, hogy rendelkezésünkre áll 11, időben közel egyforma távolságra található népszámlálási adat, két szomszédos  $i$  és  $j$  település közti kapcsolatot  $C_{ij}$ , és ezt a  $\rho_{ij}$  Pearson típusú korreláció [8,9] segítségével értelmezzük:

$$C_{ij} = |\rho_{ij}|,$$

$$\rho_{ij} = \frac{E(W_i W_j) - E(W_i)E(W_j)}{\sqrt{E(W_i^2) - E(W_i)^2} \sqrt{E(W_j^2) - E(W_j)^2}},$$

$$W_k = \frac{r_k(t+1) - r_k(t)}{r_k(t)},$$

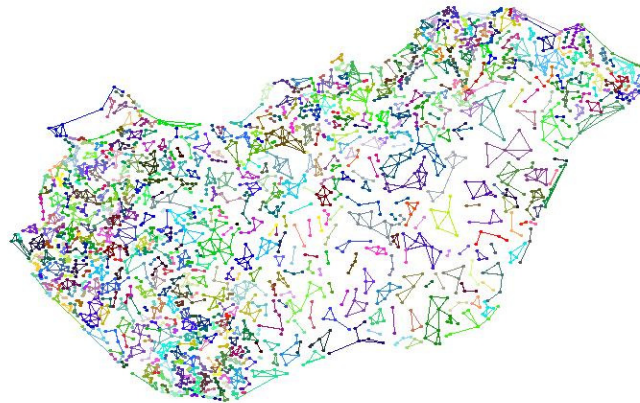
ahol  $r_k(t)$  a  $k$  település nagysága a  $t$  időpillanatban,  $E(x)$  pedig az  $x$  mennyiség átlagát jelöli. Általánosan, a Pearson típusú korreláció értéke megmutatja, hogy egy adott mennyiségnek a változása mi módon vonzza maga után egy másik ugyanolyan típusú mennyiségének változását. A Pearson típusú korrelációs együttható értéke egy 1 és -1 közti valós szám. Egy negatív együttható a mi esetünkben azt jelzi, hogy míg egyik település nagysága nő, a másiké csökken. Mivel ez is egyfajta kapcsolatot feltételez két település között, nem pedig a kapcsolat hiányát, ezért a korreláció abszolút értékével dolgozunk, elkerülve ezáltal a pozitív és negatív kötés együttes jelenléte miatt megjelenő frusztráltságot is.

A korrelációs együtthatók kiszámítása után van tehát egy lokálisan csatolt település-hálózatunk amelyen megvannak a szomszédsági kapcsolatok erőssége is. Ezen a település-hálón szeretnénk most régiókat értelmezni. A következőkben bemutatjuk a két kidolgozott, determinisztikus illetve sztochasztikus módszereket.

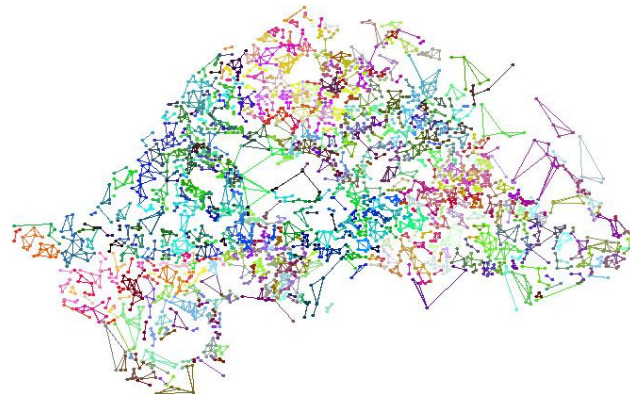
### 3.1. A determinisztikus módszer

Adva van tehát  $N$  számú település. Minden településnek ismerjük a szomszédjait, és az ezekkel való kapcsolat erősségét. Alapfeltételezésünk az, hogy minden település a vele legerősebben kapcsolódó szomszédjával egy régióba kell tartozzon. Legelső lépésen ezért, összekapcsolunk minden települést azzal a szomszédjával, akivel a legerősebb a kötése. Így kialakulnak az úgynevezett **alap**

**régió-klaszterek** (2. ábra). Az így kialakult klaszterek feltételezésünk alapján nem bonthatók szét. A következőkben, lépésenként ilyen alap klasztereket kötünk össze növelve a kialakult klaszterek nagyságát. Az egymásutáni lépések során kialakult klaszterek a további lépésekben nem bonthatók szét. A klaszterek összekötődése a következőképpen történik.



a)



b)

1. ábra, alap-régiók Erdélyre és Magyarországra

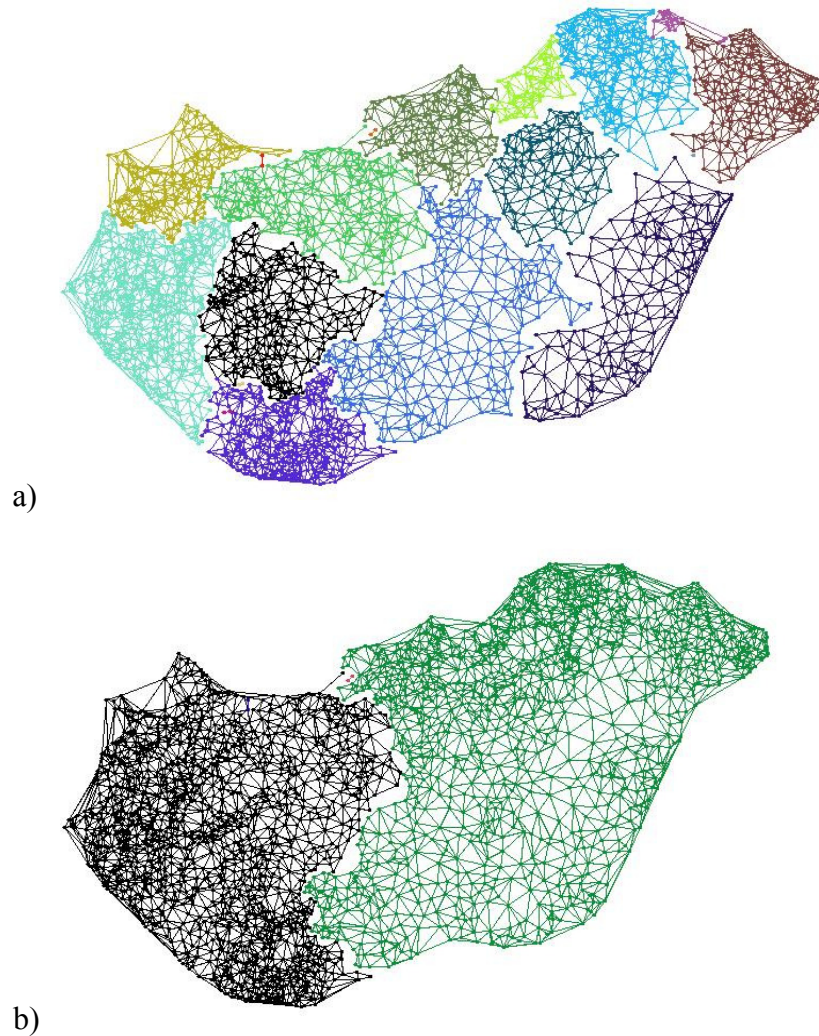
a.). minden két tetszőleges  $i$  és  $j$  klaszter között értelmezünk egy  $B_{ij}$  virtuális kötést:

$$B_{ij} = \frac{\sum_{(k \in i, m \in j)} C_{km}}{\sqrt{N_i N_j}},$$

ahol az összegzés az  $i$  és  $j$  klaszter között létező összes szomszédsági kötésre vonatkozik,  $N_k$  pedig a  $k$  klaszter nagyságát (a klaszterben levő tagtelepülések számát) jelöli, ami lényegében arányos a területével, így  $\sqrt{N_k}$  statisztikailag arányos a klaszter kerületével.



b.) Minden lépésben detektáljuk azon  $g$  és  $h$  klasztert, melyekre  $B_{gh}$  kötés maximális, majd ezeket egyesítjük. Így egyesével (3. ábra) csökkenthetjük a klaszterek számát, egészen addig, amíg el nem érjük a kívánt számú klasztert (régiót) (3., 4. ábra). Ha kezdetben volt  $K$  számú alap-klaszter,  $K-1$  lépés után minden település ugyanazon klaszterhez fog tartozni.

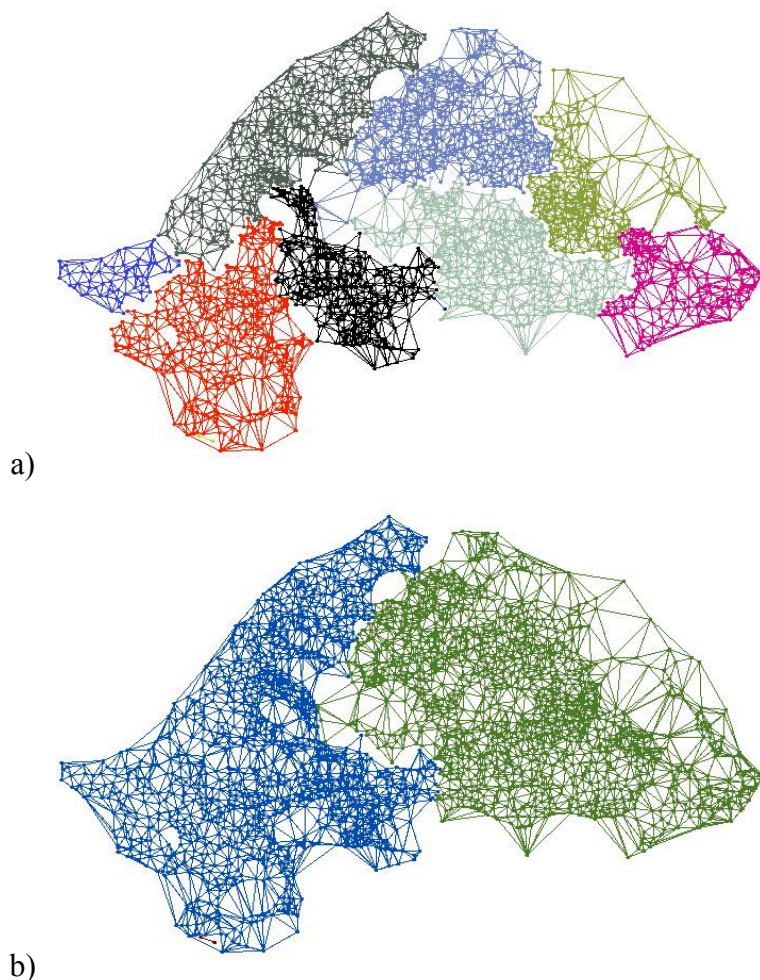


3. ábra, A determinisztikus módszerrel detektált régiók Magyarországra.

### 3.2. A sztochasztikus módszer

A sztochasztikus módszerünk, a statisztikus fizikában jól ismert Potts modell [10] Monte Carlo típusú szimulációján [11] alapszik. E modellt es módszert sikeresen alkalmaztak például a szemcsenövesztések szimulálására [12]. A Potts modell egy rácson elhelyezett skalár-változókkal jellemzett egyedek kölcsönhatását írja le. Ezen

skalár-változókat fizikus szemmel úgy is el lehet elképzelni, mint egy kvantummechanikai impulzus-nyomaték vektor (a spin is egy ilyen impulzus-nyomaték, ezért gyakran ezeket a skalár-változókat a Potts modell keretein belül spinnek nevezzük) lehetséges vetületei, ami információt ad a részecske



4. ábra, A determinisztikus módszerrel detektált régiók Erdélyre

impulzusnyomatékának az állapotáról. A Potts modell által definiált Hamilton függvény (energia operátor) a következőképpen néz ki:

$$H = -\sum_{p,q} C_{pq} \delta_{\sigma_p \sigma_q},$$

ahol a szummázás a rács elsőrendű szomszédjaira történik.  $C_{pq}$  a  $p$  és  $q$  spinek közötti kapcsolat erőssége,  $\sigma_l$  az  $l$  spin állapotát jelzi,  $\delta_{\sigma_p \sigma_q}$  pedig a Kronecker delta szimbólum, amely értéke 1, ha  $\sigma_p = \sigma_q$ , különben 0. A fenti képlet azt jelenti, hogy pozitív kapcsolat esetén ( $C > 0$ ) energetikailag az kedvező (azaz akkor kisebb a

rendszer energiája), ha két szomszédos változó hasonló értéket vesz fel, negatív kapcsolás esetén meg az, hogy két szomszédos változó különböző legyen. Adott pillanatban a rendszer energiája egyenlő az azonos állapotban levő szomszédos spinek közti kapcsolatok összegével. Az ehhez hasonló rendszerek, a hőmérséklet függvényében, egy másodfajú fázisátalakulást mutatnak. Pozitív kapcsolások esetén a kritikus hőmérséklet alatt a rendszer rendeződik, vagyis minden rácsponton levő skalár-változó hasonló értéket vesz fel. A kritikus hőmérséklet felett a rendszer rendezetlen lesz, a rácspontokon levő skalár-változók térben rendezetlenül jelentkeznek. Ha a pozitív és negatív kötések együtt jelentkeznek, a rendszer frusztrálttá válik és spinüveg-szerű nagyon komplex viselkedést mutat [13].

Az analógia az általunk értelmezett település-hálóval egyszerű. Rácsnak használjuk a már elkészített Voronoi felosztást, a rácspontoknak megfelelnek a települések, a köztük levő kötéseknek pedig a települések közötti korreláció. A korrelációt itt is modulusban használjuk, elkerülve ez által a komplex spinüveg-szerű viselkedést. A rácspontokon értelmezett skalár-változó pedig azt a régiót kódolja, amelyben az adott település van. Kezdetben mindenki külön állapotban (külön régióban) van, azaz a Potts modellre értelmezett energia zérus. Alkalmazzuk most a fizikai rendszerek azon tulajdonságát, mely szerint a rendszerek annál stabilabbak, minél kisebb az energiájuk. Ez alapján a rendszer arra törekszik tehát, hogy egy minél alacsonyabb energiájú állapotba jusson. Ez ellen dolgozik az entrópia, ami minden rendszerben egy növekvő rendezetlenséget próbál előidézni. Termodinamikai rendszerekben ez lehet egy nem-zéró hőmérséklet, szociális rendszerekben meg egyszerűen egy rendezetlenségi hatás, amelynek nagyon sok oka lehet. A rendet teremtő energia-minimalizálás és a rendezetlenséget előidéző entrópia-növekedés versengő hatása az ami ezen rendszerben dómén-struktúrát eredményez. A települések régió-dinamikáját hasonlóan képzeljük el, mint ahogyan a Potts modell változóinak időbeli változása történik egy  $T$  hőmérsékletű hőtartállyal levő kapcsolat esetén [12]. Glauber dinamikát [14] vagy BKL (kinetikus Monte-Carlo) algoritmust [15] alkalmazva a klaszter-határokon lévő településekre, egy egyszerű algoritmust képzelhetünk el, amellyel a rendszerben levő települések klaszterizálódni kezdenek. Mivel mi egy rendeződési hatást akarunk észlelni (régiók kialakulását), alacsony hőmérsékleten kell dolgoznunk, ezért célszerű a BKL szimulációs módszert használni, ugyanis a hagyományos Monte-Carlo algoritmusok esetén, alacsony

hőmérsékleten az átmeneti valószínűségek nagyon kicsik, és a rendszerben sok ideig nem történne semmi. A BKL algoritmus értelmében a települések csak a szomszédos települések állapotaiba mehetnek át, azaz egy régió-határon található település csak olyan régiókba kerülhet át, melyek tartalmazzák valamely szomszédját. Ha egy régiót (egy állapotot) egyetlen település képvisel, és ez átmegy egy másik állapotba, az adott állapot megszűnik. A BKL algoritmus alkalmazása érdekében definiálunk a rendszerben egy  $T$  hőmérsékletet (sztochasztikusságot), majd, minden egyes lépésben, a Glauber dinamika alapján kiszámítjuk a lehetséges átmeneti valószínűségeket:

$$p_n = \frac{e^{-\beta dE}}{1 + e^{-\beta dE}},$$

$$\beta = \frac{1}{k_B T},$$

$$dE = H_f - H_i,$$

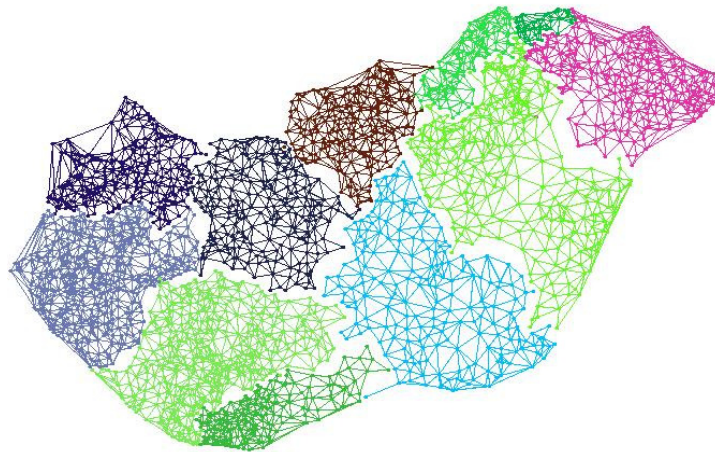
ahol  $p_n$  az  $n$ -edik átmenet valószínűsége,  $k_B$  a Boltzmann faktor,  $H_i$  és  $H_f$  pedig az átmenet előtti, valamint utáni energia. Mivel  $\beta$  csupán a  $T$  hőmérséklettől függ, általában ezt választják a szimuláció változtatható paraméterének, és nem a  $T$  hőmérsékletet. Látható, hogy az energiacsökkentő átmeneteknek nagyobb a valószínűségük, mint az energianövelő átmeneteknek. A BKL algoritmus értelmében az összes rácspontra kiszámított összes lehetséges átmeneti valószínűségek figyelembevételével felépítünk egy kumulatív sort:

$$P_n = P_{n-1} + p_n, P_0 = 0.$$

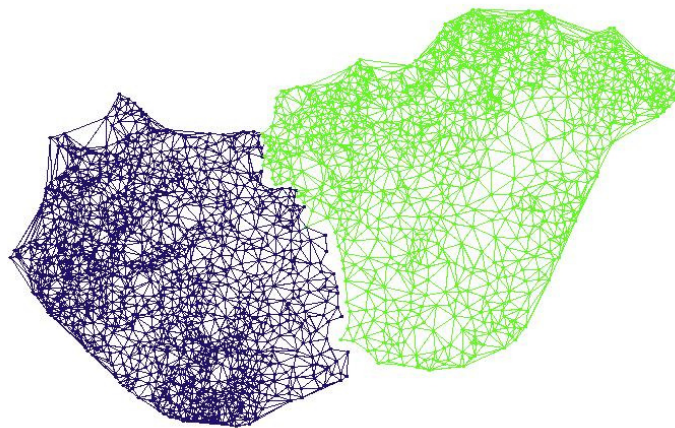
Véletlenszerűen kiválasztunk egy  $r \in [0, P_{n_{max}}]$  számot, megkeressük az  $x$ -edik átmenetet, amelyre igaz, hogy  $P_{x-1} < r \leq P_x$ , és végrehajtjuk ezt. Ezt a lépést több milliószor alkalmazva a rendszerben egy klasszikus Monte Carlo típusú algoritmust értelmezünk. Ha a rendszerben  $N$  spinünk van, akkor  $N$  számú átmenetet tekintünk egy Monte-Carlo lépésnek.

A rendszer egy bizonyos termalizációs idő után beáll egy viszonylag stabil állapotba. A definiált hőmérséklettől függően kialakulnak kisebb-nagyobb klaszterek, melyeknek határvonalai fluktuálnak (5., 6. ábra). Kis hőmérsékletek (azaz nagy  $\beta$ ) esetén nagyon hamar a rend dominál és a doméniumok (régiók) erősen megnőnek, míg a magas hőmérsékletek esetén a doméniumok nagyon kicsik lesznek. A kialakuló domén-szerkezet azonban nemcsak a hőmérséklettől függ, hanem attól is, hogy milyen sok ideig alkalmazzuk a BKL dinamikát (azaz függ a  $K$ , Monte-Carlo lépések számától). Úgy a hőmérséklettel, mint pedig a szimulációs idővel is játszódhatunk,

ahhoz hogy megfelelő számú es méretű régiót alkossunk. A fluktuáló dómén-határok is értelmezhetők. Az egyik doméniumból a másikba „lépegető” települések valószínű, hogy mindkét régióhoz hozzátartoznak (azaz a régió-határ nem egy konkrét vonal, hanem egy sáv, ahogy az valószínűleg a valóságban is van). Ezt a jelenséget is ki lehet használni arra, hogy a régiókat meg jobban értelmezzük.



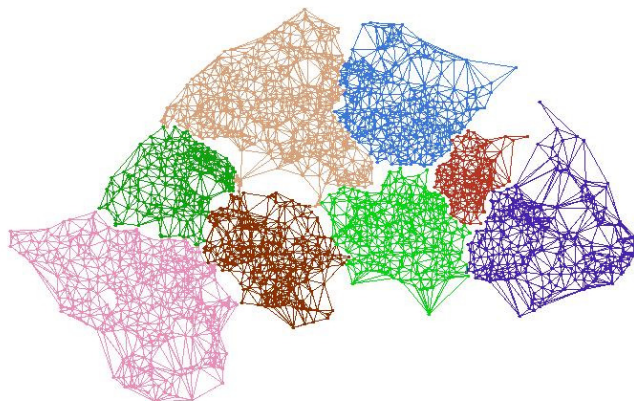
a)  $\beta=1.0, K=1000$



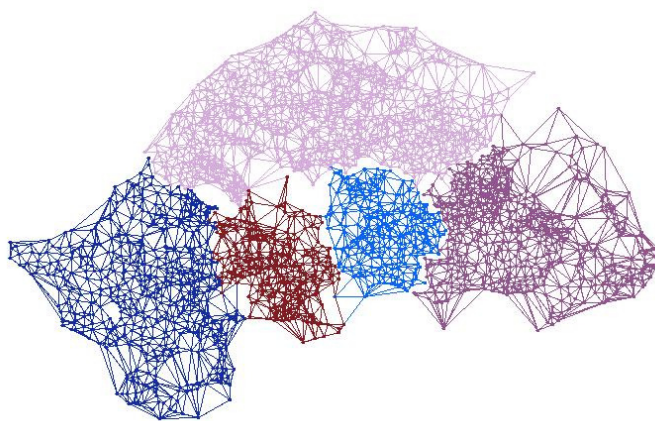
b)  $\beta=1.0, K=2500$

5. ábra, A sztochasztikus módszer eredményei Magyarországra.





a)  $\beta=1.0, K=800$



b)  $\beta=1.0, K=1700$

6. ábra, A sztochasztikus módszer eredményei Erdélyre

## 4. Tárgyalás és Következtetések

Amint az ábrákon is látható, mindkét módszer elfogadható régió-szerű klasztereket határol le. Ezek nem tagolt szélűek, nem túlságosan elnyúlóak. Mindemellett valamilyen szinten felismerhetők a különböző, akár történelmi régiók is, bár ezeknek határai elmosódtak, átalakultak. Ez nyilván nem meglepő, hiszen az elmúlt évszázadban rengeteg átrendező-hatás érintette a vizsgált területeket, kezdve a szocialista rezsim működési elveitől, egészen akár a globalizációig. Érdeemes megemlíteni Magyarország esetén, hogy a Dunának kiemelt szerepe és jelentősége van ezen régió-klaszterek alakulásában. Mindkét módszer esetében a folyó látható módon, stabilan kettéosztja az ország területét. Ha kisebb skálán vizsgálódunk, felismerhetünk akár olyan területeket, mint Észak- és Dél-Alföld, vagy Közép-

Nyugat- és Dél-Dunántúl. Erdély esetében elsősorban egy kelet és egy nyugat Erdélyt látunk. Kelet Erdély jól egyezik a mindenki által ismert történelmi Erdéllyel, leszámítva ennek dél-nyugati részét, ami szakemberek szerint a valóságban is inkább Temesvár központúvá vált. Kisebb skálán itt is szépen elkülönülnek olyan területek, mint a Bánát, Székelyföld, Szászvidék, Észak-Erdély vagy Partium. Nyilván az országok jelenlegi közigazgatási felosztását, azaz a megyéket nem kaphatjuk vissza, hiszen ezek határait többnyire politikai megfontolások alapján határozták meg, gyakran figyelmen kívül hagyva, a települések kapcsolatát sokszor jól jellemző kulturális, szociális, demográfiai, de akár gazdasági szempontokat is. Hogy a jelenlegi helyzet talán nem a legkedvezőbb az országok gazdasági fejlődésének szempontjából, azt módszereink is alátámaszthatják, nyilván sok más adattípust kéne még vizsgálni, és egyéb szempontokat is figyelembe venni, ahhoz, hogy ezt biztosan állíthassuk.

Manapság a politikában is egyre nagyobb hangsúlyt fektetnek a régiók pontos meghatározására, mivel ezek nagyban befolyásolják egy adott terület gazdasági hatékonyságát, de nem elhanyagolhatóak kulturális szempontból sem. Európai Uniós törekvések szerint egy decentralizációs folyamat megy végbe, amely keretén belül nem egyes országoknak utalják át a költségvetésből a megítélt összegeket és támogatásokat (hogy aztán ezek tetszésük szerint osszák szét), hanem egyenesen az erre a célra kialakított fejlesztési, úgynevezett makro-régióknak. Ez egy aktuális probléma a mai Romániában, és kiemelt jelentőségű, hogy ezen makro-régiókat (melyek már léteznek, de közgazdászok szerint „összetételük” nem helyes) pontosítsuk, ugyanis sok múlhat ezen, az ország gazdasági fejlődését tekintve, mely kihatással van a saját jövőnkre (azaz egy jól működő rendszer emelheti az életszínvonalat). Evidens, hogy ezek meghatározásához, vagy pontosításához sem elegendő csupán egy adott típusú adatot figyelembe venni, de amint azt dolgozatomban már megjegyeztem, a módszerek, bizonyos feltételek mellett működnek egyéb jellegű adatokkal is. Nyilván mi a legkézenfekvőbbet, és a bárki számára legkönnyebben hozzáférhetőbbet használtuk.

Fontosnak tartom kiemelni, hogy módszerünk erőssége nem a feldolgozott adatok típusában van, hanem azon, hogy mellőz bármilyen szubjektivitást, legyen ez egyéni megfontoláson alapuló, de akár nemzeti, vagy politikai érdekekből fakadó is. Nem azt állítjuk, hogy egy precíz módszert szolgáltatunk, ami százszázalékosan jó eredményhez vezet, de mindenképpen egy olyan alkalmazása a matematikának, fizikának, földrajznak, ami sok, régiókkal foglalkozó szakember, legyen az

földrajztudós, közgazdász, szociológus, vagy akár politikus, kezében egy olyan eszköz, amely sok esetben segítséget nyújthat, és rávilágíthat egy objektív megoldásra.

## 5. Hivatkozások

- [1] Charlie Karlsson, *Clusters, Functional Regions and Cluster Policies*,  
<http://www.infra.kth.se/cesis/documents/WP84.pdf>
- [2] Charlie Karlsson, Michael Olsson, *The identification of functional region: theory, methods, and applications*, Ann Reg Sci 40:1 – 18, 2006
- [3] William Mitchell, Anthea Bill and Martin Watts, *Identifying functional regions in Australia using hierarchical aggregation techniques*,  
<http://e1.newcastle.edu.au/coffee/pubs/wp/2007/07-06.pdf>
- [4] Varga E. Árpád, *Erdély etnikai és felekezeti statisztikája*,  
<http://varga.adatbank.transindex.ro/index.php3>
- [5] ASTROFORUM - Locul de intalnire a pasionatilor de astrologie si nu numai  
<http://www.astroforum.ro/index.php?showtopic=90>
- [6] Járai-Szabó Ferenc and Zoltán Néda, *On the size distribution of Poisson Voronoi cells*, Physica A, Volume 385, Issue 2, 15 November 2007, Pages 518-526
- [7] F. Jarai-Szabo, Z. Neda, *On the size-distribution of Poisson Voronoi cells*  
<http://arxiv.org/abs/cond-mat/0406116>
- [8] Values of the Pearson Correlation  
<http://cnx.org/content/m10950/latest/>
- [9] Correlation, From Wikipedia, the free encyclopedia  
<http://en.wikipedia.org/wiki/Correlation>
- [10] F. Y. Wu, *The Potts model*, Rev. Mod. Phys. 54, 235 - 268 (1982)
- [11] Wikipedia, Monte Carlo method  
[http://en.wikipedia.org/wiki/Monte\\_Carlo\\_method](http://en.wikipedia.org/wiki/Monte_Carlo_method)
- [12] D. Weygand, Y. Brechet and Z. Néda; "Capillarity-driven interface dynamics: application to grain growth phenomenon", Phil. Mag. B, vol. 75, 937 (1997)
- [13] Néda, Z., Florian, R., Ravasz, M., Libál, A., Györgyi, G., *Phase transition in an optimal clusterization model*. — Physica A 362: 357-369 (2006)



[14] David A. Levin, Malwina J. Luczak, Yuval Peres, *Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability*

[http://arxiv.org/PS\\_cache/arxiv/pdf/0712/0712.0790v2.pdf](http://arxiv.org/PS_cache/arxiv/pdf/0712/0712.0790v2.pdf)

[15] Arthur F. Voter, *Introduction to the Kinetic Monte Carlo Method*

[http://www.ipam.ucla.edu/publications/matut/matut\\_5898\\_preprint.pdf](http://www.ipam.ucla.edu/publications/matut/matut_5898_preprint.pdf)