



A gravitációs modell kalibrálásának alapkérdései

The basic questions of the calibration of gravity model

Dusek Tamás

Széchenyi István Egyetem

E-mail:

dusekt@sze.hu

A gravitációs modell jól ismert a magyar kutatók előtt. Hazai alkalmazásai azonban szinte kivétel nélkül elméleti vonzáskörzetek és elméleti áramlásnagyságok meghatározására szorítkoznak, és nem foglalkoznak a modell ismert áramlási adatok melletti kalibrálásával. Jelen tanulmány a kalibrálás alapkérdéseit mutatja be, egy szemléltető példa segítségével és korábbi nemzetközi eredmények ismertetésével, remélhetőleg ösztönzésül szolgálva a modell tapasztalati tudomány szellemében történő jövőbeli alkalmazásához.

Kulcsszavak:

gravitációs modell,
migráció,
területi statisztika,
regressziószámítás

Keywords:
gravity model,
migration,
spatial statistics,
regression analysis

Gravity model is a well-known method for investigating spatial interactions. However, its applications in Hungary deal almost exclusively with the delimitation of theoretical catchment areas and theoretical flow sizes and neglect the calibration of model parameters using empirical flow data. The aim of this paper is to show the basic questions of calibration using an illustrative example and describing previous international results, which hopefully give encouragement to the future application of the model in the spirit of empirical science.

Beküldve: 2016. június 1.

Elfogadva: 2016. június 30.

Bevezetés: a gravitációs modell kettős használata a tapasztalati adatok szempontjából

A gravitációs modell jól ismert a magyar területi kutatók előtt, elméleti leírására és alkalmazott kutatásokban történő használatára is számos példát lehet találni. Mindazonáltal a modell hazai alkalmazásai nagyrészt kimerülnek a nem ismert területközi áramlások becslésében, egyes központok elméleti vonzáskörzeteinek lehatárolásában, valamint elméleti potenciálmezők feltárásában. Ezen kutatások között találunk európai léptékűt (Tagai 2011, Tóth–Kincses 2007), országos kiterjedésűt (Lackó 1978, Bajmócy–Kiss 1999, Nagy 2011, Kincses–Tóth 2012, Kincses–Tóth 2014), regionálisat (Kiss–Bajmócy 2001, Péntes 2005, Tóth 2005, Hardi–Szörényiné 2014), egyetlen településre vonatkozót (Beluszky 1974, Papp 1981, Sziva 1987, Bella 2011) és településen belüli áramlásokra vonatkozót (Nagy 1996). Mindegyik esetben valamilyen tömegdefiníciót (leggyakrabban a népességszámot) és távolságmeghatározást alkalmaznak, önkényesen megválasztva a távolságfüggést befolyásoló kitevő nagyságát. A modell kalibrálásával¹ a külkereskedelem elemzése kapcsán találkozhatunk (Udvari 2012, Neumanné 2014).

A modell kalibrálás nélküli alkalmazásait inkább annak a kényszerű helyzetnek lehet tulajdonítani, hogy a területközi áramlásokra vonatkozó adatok rendkívül szűkösek. Így a modell által meghatározott elméleti vonzáskörzetek és elméleti áramlásnagyságok tényleges, tapasztalati adatokkal való összevetése vagy egyáltalán nem lehetsé-

¹ A kalibrálás kifejezés a paraméterek becslése, a modell specifikációja szinonimájaként használatos.

ges, vagy csak korlátozottan, inkább az ösztönös megérzésre, mintsem konkrét áramlási adatokra alapozottan. Mindemellett az eredmények a tényleges helyzettől (vagyis az áramlások tényleges megfigyelése révén kapott befolyási zónáktól) nagyon nem tudnak eltérni. A közúti elérhetőségeken alapuló körzetlehatárolásokhoz képest (például Szalkai 2012) pedig annyi lesz a különbség, hogy a gravitációs modellel lehatárolva a nagyobb népességű (súlyú) központoknak valamivel nagyobb, a kisebbeknek kisebb lesz a vonzáskörzete. Emellett a gravitációs modellt használva a nagy népességszámú központokhoz a térbeli folytonosság követelményét megszegő módon exklávék is tartozhatnak.

A modell ilyen célú felhasználása ugyanakkor nagyon különbözik a fizikai analógiájának tekintett gravitációs tömegvonzás alaptörvényétől. Utóbbi esetben megengedhetetlen lenne önkényes módon meghatározni a távolsághatást kifejező kitevőt: ez a tapasztalati megfigyelésekkel összhangban kettő lesz és nem akármekkora. A kitevővel kapcsolatos manipulációk veszélyét egy analógiával lehet leginkább megvilágítani. Képzeljük el, hogy egy kutató a jövedelmek és a megtakarítások kapcsolatát vizsgálja. Nagyobb jövedelem mellett nagyobb megtakarítás várható, ezt adatok nélkül is tudjuk. A jövedelem és a megtakarítás közötti kapcsolatot leíró konkrét regressziós függvényt azonban csak adatok, tapasztalati megfigyelések révén fejezhetjük ki. A regressziós függvény az adott helyre és időre vonatkozó megfigyelésekre lesz érvényes, és alkalmas időbeli, térbeli, a sokaság alcsoportjai közötti összehasonlításokra (például x városban, régióban, nemzetnél stb. nagyobb a megtakarítási ráta, mint y városban, régióban, nemzetnél). Adatok hiányában a regresszió paraméterei ugyanolyan hasraütésszerűen határozhatók meg, mint a gravitációs modell távolságkitevője. Adatok nélkül a kitevő nagyságára nem adható olyan becslés, amit ellenőrizni lehetne, ezért a modell leíró célú jószágáról, előrejelzésre és tervezésre való felhasználhatóságáról semmit sem lehet megállapítani.

A modell kétféle, tapasztalati adatok nélküli és tapasztalati adatokon alapuló használatát a következőkben lehet még szemléletesebben megfogalmazni. Első esetben nincsenek ismert áramlási adatok, azt a modell szolgáltatja:

*Áramlás modell szerint = konstans * kiindulópont tömege * végpont tömege / távolság*

A második esetben rendelkezésre állnak az áramlási adatok:

*Áramlás ténylegesen = (konstans * kiindulópont tömege * végpont tömege / távolság) * hiba*

A tömeget és a távolságot mindkét esetben konkrétan kell meghatározni. Második esetben a cél a tényleges áramlások minél pontosabb, minél kisebb hibával történő reprodukciója. Általánosságban minél kisebb a hiba, a térbeliség szerepe annál jelentősebb, minél nagyobb a hiba, a térbeliség, a távolság annál kevésbé gyakorol befolyást az áramlások nagyságára. Nagy hiba mellett a távolság és a tömeg eltérő definiálása is segíthet a becslés pontosításában. Ha a távolság esetleg nem növeli a magyarázó erőt, akkor az adott jelenség áramlását nem befolyásolja a földrajzi elhelyezkedés. Egy ilyen tapasztalati eredmény meglehetősen furcsa és kivételes lenne.

Jelen tanulmányban a gravitációs modell kalibrálásával kapcsolatos általánosabb kérdéseket tekintjük át, nemcsak absztrakt módon, hanem egy konkrét és önmagában is elemzésre érdekes illusztratív példa segítségével. Ez a módszer második típusú használatát jelenti, aminek tehát az az előfeltétele, hogy egy mátrix formájában rendelkezésre álljanak az áramlási adatok. A rendkívül gazdag részletesebb elemzési lehetőségek tárgyalása egy későbbi tanulmány tárgyát képezi.

Az áramlási adatok szűkössége

A területi adatok nagy része egyetlen területegységet vagy pontot ír le, térbeli értelemben statikus. Az egynél több területegységre vonatkozó, térbeli kapcsolatokat leíró és területközi áramlásokra vonatkozó adatok azok fontosságához képest rendkívül szűkösek. Az első esetben az áramlási adatok vagy egy pont és az összes többi pont közötti áramlásokra vonatkoznak (például egy településre más településekből ingázók, egy országba más országokból érkező külföldi turisták), vagyis egy sorban vagy oszlopban megadhatók. A második esetben az összes pont összes többi pont közötti áramlására vonatkozóan ismerjük az áramlási adatokat, amelyek ekkor egy $n \times n$ méretű (n a területegységek száma) mátrix formájában adhatók meg. Az utóbbi esetből mindig előállítható az első, egyszerűbb eset, ehhez csak egy területegységet kell kiválasztani, és a teljes áramlási mátrixnak csak egy sorát vagy oszlopát, az oda- vagy eláramlásokat figyelembe venni. Az első eset ugyanakkor nem feltétlenül bővíthető, nemcsak a gyakorlati adathiány miatt, hanem az adatok elvi hiánya miatt is, amennyiben csak egy vagy néhány vonzópont van és sok vonzott, amelyek között az áramlások nem értelmezettek. Például, az egy egyetemre beiratkozott hallgatók megadhatók lakóhely szerint, és akkor az első esethez jutunk, de a hallgatók lakóhelye közötti távolság és áramlás nem értelmezett, legalábbis úgy biztosan nem, mint az egyetem és a lakóhely közötti távolság.

Már az első esetben is nagyméretű adattáblákra lehet szükség, a másodikban pedig hatványozódik az adatoméret. Az adatok szűkössége ezzel, vagyis az adatközlés helyigényével és a megfigyelhetőség gyakorlati nehézségeivel függ össze. Például 7 régió közötti áramlások megadásához egy 7×7 -es mátrixra van szükség (49 cella), ami még átlátható méret a hagyományos papíralapú adatközlés számára is, részben akár szabad szemmel is feldolgozható. 20 megye között ugyanakkor már $20 \times 20 = 400$, 3000 település között $3000 \times 3000 = 9$ millió cellából álló mátrixra lenne szükség, a nagyjából 600 ezer indiai települést tartalmazó mátrix pedig 360 milliárd cellából állna. Ebben a tanulmányban a teljes áramlási mátrixra vonatkozó vizsgálatokat tekintjük át. Ilyenkor a vizsgált terület egészét egyetlen egyenlettel jellemzik. A területegységenkénti elemzési lehetőségeket egy későbbi tanulmányban tárgyaljuk majd, amikor is a mátrix sorait és oszlopait külön-külön kezeljük, így lehetőségünk lesz területegységenként és irányonként eltérő távolsághatások számítására.

Olykor (például közlekedési, utazási adatoknál) csak az összes beáramlás és az összes kiáramlás ismert, vagyis a mátrix összegző sorai. Ez a mindkét oldalról korlátos (vagy korlátozott) gravitációs modell megnevezés alatt tárgyalt eset a teljes mátrix ismeretéhez képest teljesen eltérő problémákat vet fel. Ilyenkor becsléseket kívánnak adni a mátrix belső celláinak értékére, a kiindulási és a végpontok közötti áramlások nem ismert nagyságaira. Ezzel a kérdéssel sem foglalkozunk.

1. táblázat

Példák a lehetséges áramlási adatokra

Examples of potential flow data

Alany	Áramlás
Emberek	Migráció
	Ingázás (munka, iskola)
	Turizmus
	Vásárlás, szórakozás
Áruk	Régióközi kereskedelem
	Szállítási teljesítmény
Információ	Telefonhívások

Az áramlások alanyai lehetnek például emberek, termékek, energia, pénz, betegségek, eszmék, ötletek. Bármilyen változás, ami történik körülöttünk a világban, nagyrészt e tényezők áramlása révén valósul meg. Az 1. táblázat néhány lehetséges áramlási adatot tartalmaz a társadalomföldrajz területéről. Ezek közül a migrációra, a turizmusra és a külkereskedelemre vonatkozó mátrixok állíthatók össze a hagyományos adatközlésben elérhető adatokból. A többi adat egyedi felméréssel vagy meglévő adatokból (például népszámlálási ingázási információk) egyéni adatlekéréssel állítható elő.

Kalibrálás regressziószámítással

Az áramlási mátrix ismeretében négy változóval rendelkezünk. Ezek jelöléseit a 2. táblázatban soroltuk fel (a mai napig nem alakult ki olyan egységes jelölésük, mint például a regressziószámításnál). Az áramlási adatok, a kiinduló és a végpontok, valamint a köztük lévő távolságok ismeretében az elsődleges cél egy olyan egyenlet megadása, amelynek segítségével a tényleges áramlások a lehető legpontosabban, máshogyan kifejezve a lehető legkisebb hibával reprodukálhatók. A kiinduló és a végpontok, valamint a köztük lévő távolság közötti szorzatszerű (multiplikatív) kapcsolatot feltételezve a többváltozós hatványkitevős regresszió használható fel ennek megoldásához. Ehhez elegendő lenne a három magyarázó tényező (kiindulópont tömege, végpont tömege, távolság) közötti szorzatszerű kapcsolat feltételezése, nem szükséges a tényezők előjelét, a kapcsolat irányát előre rögzíteni. Mindazonáltal a hétköznapi ta-

pasztaltnak, ösztönös megérzésnek, gondolkodásnak teljes mértékben megfelel a tömeg növekedésével nagyobb, a távolság növekedésével kisebb kapcsolatot feltételezni, ezért az egyenlet felírásában érvényesítettük ezt a szokásos feltevést, vagyis a tömegeket a számlálóban, a távolságot a nevezőben helyeztük el. A korábbi általános felírásban a tömegekhez nem tartoztak kitevők, máshogyan fogalmazva, a kitevő értékét egynek feltételeztük. Erre a megszorító feltevésre nincs szükség, az egységnyi rugalmasságtól való eltérés lehetősége az összefüggés még általánosabb formáját jelenti.

2. táblázat

A gravitációs modell kalibrálása regresszióval, jelölések és egyenletek

Calibration of gravity model with regression: notations and equations

Megnevezés	Jelölés, egyenlet
Általános modell	$\hat{y}_{ij} = k \times \frac{P_i^a P_j^b}{d_{ij}^c}$
Linearizált általános modell	$\log_{10} y_{ij} = \log_{10} k + a \times \log_{10} P_i + b \times \log_{10} P_j - c \times \log_{10} d_{ij}$
Az i-edik és j-edik területegység közötti áramlás nagysága	y_{ij}
Az i-edik területegység tömege	P_i
A j-edik területegység tömege	P_j
Az i-edik és a j-edik területegység közötti távolság	d_{ij}
Konstans	k
A tömeghez és a távolsághoz tartozó kitevők	a, b, c
A területegységek száma	n
A területegységek közötti kapcsolatok száma	$n*(n-1)$

A 2. táblázat nem tartalmazza a regressziószámítás ismert egyenleteit, például az együttthatók meghatározását vagy a determinációs hányados számítását és annak felbontását. A hatványkitevős regresszió a változók logaritmizálásával visszavezethető a lineáris regresszióra, legalábbis a számítások leíró statisztikai részét tekintve. Mindössze annyi a feladat, hogy mind a négy (áramlás, két tömeg és távolság) változót logaritmizálni kell, majd ezután a logaritmizált változókkal kell elvégezni a lineáris regressziószámítást. Az áramlási adatok ismeretében így egyszerre ellenőrizhető az áramlásokra vonatkozó általános feltevés (a távolsághatás létezése) és többféleképpen (kitevő, determinációs együtttható, standardizált kitevő) meghatározható annak konkrét nagysága is.

A három magyarázó változós számítási mód mellett létezik a változók bizonyos előzetes transzformálását igénylő két magyarázó változós és egy magyarázó változós módszer is. A két magyarázó változósnál a kiinduló és a végpont tömegét megszorozva egy új változót képeznek, ami ekkor a tömeghatást fejezi ki, mellette pedig meg-

marad természetesen másik magyarázó változónak a távolság. Az egyenlet ekkor a következő:

$$\hat{y}_{ij} = k \times \frac{(P_i \times P_j)^a}{d_{ij}^b}$$

A logaritmizálást követő lineáris forma pedig:

$$\log_{10} \hat{y}_{ij} = \log_{10} k + a \times \log_{10} (P_i \times P_j) - b \times \log_{10} d_{ij}$$

Az egy magyarázó változós formánál előzetesen az áramlási adatokat osztják el a kiinduló és a végpont tömegének a szorzatával. Ezzel tömegarányos áramlást kapnak, amelynek egyetlen magyarázó változója a távolság lesz:

$$\frac{\hat{y}_{ij}}{P_i \times P_j} = k \times \frac{1}{d_{ij}^a}$$

A logaritmizálást követő lineáris forma:

$$\log_{10} \frac{\hat{y}_{ij}}{P_i \times P_j} = \log_{10} k - a \times \log_{10} d_{ij}$$

Ezt a változatot mutatja be Illés Iván (1975), valamint Nemes Nagy József és Piros György (1984). Az egy magyarázó változós forma alkalmazását elvileg egyrészt az indokolná, hogy szinte mindig elsősorban a távolsághatás az érdekesebb és nem a tömeg szerepe. Másrészt az egy magyarázó változóhoz előkészített formában felírt adatokat egy kétdimenziós koordináta-rendszerben lehet ábrázolni. Ugyanakkor a két magyarázó változósban a tömegekhez egyetlen közös kitevő tartozik, az egy magyarázó változósban pedig tovább csökken a tömeg szerepe, ugyanis azt csak az adatok tömeg szerinti fajlagosításánál használjuk, direkt módon nem vonjuk be a regresszióba. Így önként eltekintünk már meglévő információktól, feleslegesen és mindenképpen csökkentjük a becslés jóságát. A három változat számításigénye hasonló, így ez sem indokolja a változók számának csökkentését. Történetileg a legelterjedtebb forma a három magyarázó változós, ez gyakorlati érv az előnyben részesítése mellett, de ennél fontosabb, hogy elvileg is ez a leginkább alátámasztható.

A kalibrálás alaperedményei a megyék közötti magyarországi migráció esetében

Az eddigiekben absztrakt módon tárgyaltuk a gravitációs modell kalibrálásának kérdését. A következő példával pontosabban megmutathatjuk a kalibrálásban rejlő lehetőségeket. A magyarországi megyék közötti migrációra vonatkozó adatokat használjuk fel, mert azok könnyen beszerezhető és jól értelmezhető térfelosztásra vonatkozó általános adatok. Összesen 14 év, a 2000-től 2013-ig terjedő időszak adatait használjuk fel. A több év használata az eredmények stabilitásának ellenőrzése szempontjából előnyös. 2013-ra vonatkozóan részletesebb elemzési lehetőségeket is bemutatunk.

A területi egységek tehát a megyék lesznek, amelyek egyértelműen értelmezhetők. A közöttük lévő távolság meghatározására ugyanakkor több megoldás is elképzelhető,

amelyek közül a gyakorlatban kettő a leginkább elterjedt: a megyeszékhelyek közötti távolságok, illetve a megyék súlypontjai közötti távolságok. Mindkét változatnak több esete különböztethető meg: a megyeszékhelyek közötti távolság értelmezhető például légvonalban, közúton, időben vagy valamilyen költségben, a megyék súlypontjai pedig szintén többféleképpen határozhatók meg. Alapesetben a népesség súlypontjaként választjuk ki meg a megyék központját, a közöttük lévő távolságot pedig az e súlypontok közötti légvonalbeli távolságként. A megyeszékhelyek közötti távolság alkalmazása akkor lenne indokolt, ha csak a megyeszékhelyek közötti áramlásokat vizsgálnánk.

A kalibrálás eredményeit a 3. táblázatban mutatjuk be. A távolsághatás kitevője mindig negatív, a szokásnak megfelelően az abszolút értékét adjuk meg a későbbiekben is. A népességszám évről évre kismértékben változik, a súlypontok is elmozdulnak, így évről évre más távolságokkal kellene számolnunk. Az elmozdulás azonban olyan rendkívül kicsi, hogy a 2000-es és a 2013-as távolságok (vagyis a népesség adott évi megyei súlypontjai közötti légvonalbeli távolságok) közötti regresszióhoz tartozó determinációs együttható értéke 99,98%. Ezért végig a 2013-as adatokkal számoltunk. A népesség arányeltolódása kicsit nagyobb, ezért évről évre az adott év népességszámát használtuk. A 14 év eredményeit összehasonlítva megállapítható, hogy egymáshoz nagyon hasonlóak az eredmények, ami a migrációs tendenciák hasonlóságának is következménye (bár ugyanolyan kitevők egészen más migrációs irányok mellett is lehetségesek).

3. táblázat

A távolsághoz tartozó kitevő abszolút értékének alakulása

Change in the absolut value of distance exponent

Év	Kitevő nagysága			Determinációs együttható, %
	távolság (abszolút érték)	kiindulópont	végpont	
2000	1,40	1,17	1,03	86,5
2001	1,38	1,13	1,04	86,6
2002	1,35	1,13	1,13	85,8
2003	1,38	1,13	1,12	86,4
2004	1,35	1,09	1,14	85,9
2005	1,36	1,08	1,16	86,1
2006	1,38	1,09	1,23	85,4
2007	1,39	1,09	1,26	83,6
2008	1,39	1,12	1,24	82,9
2009	1,38	1,08	1,26	85,7
2010	1,33	1,11	1,24	85,3
2011	1,31	1,09	1,20	85,4
2012	1,32	1,09	1,17	86,3
2013	1,30	1,07	1,19	85,8

A 2013-ra vonatkozó egyenlet a következő:

$$\hat{y}_{ij} = 10^{-7,912} \times \frac{P_i^{1,071} \times P_j^{1,193}}{d_{ij}^{1,304}}$$

Az egyenlet értelmezése a következő. Ha két megye között 1%-kal növekszik a távolság, akkor 1,304%-kal csökken a migráció nagysága (*ceteris paribus*). Ha a kiinduló megye népessége 1%-kal növekszik, akkor a migráció 1,071%-kal, ha a végpont megye népessége 1%-kal növekszik, a migráció 1,193%-kal. A migráció arányait tekintve tehát kicsit nagyobb mértékben növekszik a magasabb lélekszámú megyék körében. Ez valószínűleg elsősorban Budapest és Pest megye együttes hatásával függ össze. Budapest és Pest megye elhagyásával a végponthoz tartozó kitevő 1,012-re csökkenne (gyakorlatilag 1-re), a kiindulópontához tartozó viszont 1,237-re nőne. Ebben jelentős szerepe van annak, hogy a magas lélekszámú Szabolcs-Szatmár-Bereg és Borsod-Abaúj-Zemplén megyében az átlagosnál nagyobb az elvándorlás aránya. Mindenesetre a kitevők 1-hez közeli értéke alapján a nagyjából népességarányos áramlás feltevése is elfogadható, amellet, hogy a gyakorlatban megfigyelhető, ettől eltérő ingadozások egyedi magyarázatot igényelnek.

A fenti egyenletbe behelyettesítve a becsült áramlásokat kaptuk. Például, Fejér és Zala megye távolsága 134,4 kilométer, Fejér megye népessége 432 559, Zala megye népessége 308 837 fő. Ezt behelyettesítve:

$$\hat{y}_{ij} = 10^{-7,912} \times \frac{432559^{1,071} \times 308837^{1,193}}{134,4^{1,304}} = 79,9$$

vagyis a modell által meghatározott áramlás 80, míg a tényleges áramlás 90 fő volt.

A determinációs együttható 85,8%, és évről évre nagyon stabilan hasonló értékeket mutat. A magas érték rendkívül figyelemreméltó, hiszen egyetlen tulajdonságra, viselkedésre vonatkozó hagyományos magyarázó változó sem szerepel a független változók közül, csupán a megyék népessége és távolsága.

A szorzatszerű összefüggés feltételezése azon alapult, hogy két csoport elemei között a lehetséges interakciók száma az elemek szorzatával és nem az összegével egyenlő. Az adatok egyéb típusú kapcsolatokkal történő összehasonlításra is alkalmazhatók. Ha lineáris kapcsolatot feltételezünk, vagyis a távolság, a kiinduló népesség és a végpont népesség hatása összeadódna, akkor az ehhez a regresszióhoz tartozó determinációs együttható mindössze 26,2% lenne (2013-ban). Így ha a két megye közötti távolság egy kilométerrel növekszik, akkor a migráció nagysága 1,76 fővel csökken, ami szintén az elvártnak megfelelő előjelű eredmény, de sokkal kisebb magyarázó erő, ezzel együtt nagyobb hiba mellett.

A három tényező külön-külön hozzájárulását a determinációs együtthatóhoz annak vizsgálatával lehet mérni, hogy mennyivel növekedne a determinációs együttható, ha az adott változót utolsónak vonnánk be az egyenletbe. Ezeket a számításokat kell elvégezni a multikollinearitás Theil-féle mutatójának a meghatározásához is. Emellett megvizsgálhatjuk azt is, hogyha csak az egyik magyarázó változó szerepelne, akkor az

önmagában mekkora determinációs együtthatóval járna együtt. A 4. táblázat azt mutatja, hogy önmagában a távolság szerepe a legnagyobb, a multikollinearitás nagyon kicsi. Jelen esetben a multikollinearitás egyik oka, hogy Budapest és Pest megye, a két legnagyobb népességű terület egység közel van egymáshoz. Ezt leszámítva nincs további jelentős kapcsolat a területi elhelyezkedés és a népességszám között.

4. táblázat

Az egyes változók hozzájárulása a determinációs együtthatóhoz

Contributions to the coefficient of determination by variables

	(%)
Változó	Hozzájárulás
Távolság	31,9
Kiindulópont népessége	20,3
Végpont népessége	25,1
Egyedi hozzájárulások összesen	77,3
Közös rész (multikollinearitás)	8,5
Egyes változók magyarázó ereje önmagában, egy magyarázó változós regresszió esetén	
Távolság	43,0
Kiindulópont népessége	22,8
Végpont népessége	28,3

A három tényező relatív fontosságát a standardizált együtthatók alapján is megítélhetjük. Nagyságrendnyi különbség nincs közöttük, de ez alapján is minden évben rendre a távolság szerepe lesz a legnagyobb. Például 2013-ban a távolsághoz tartozó standardizált együttható abszolút értéke 0,569, a végpont népességé 0,504, a kiinduló népességé 0,453 volt.

Ha a megyék közötti távolságot nem a népesség súlypontjai, hanem a megyeszékhelyek közötti közúti távolságok alapján értelmezzük, akkor az így módosított regresszió (2013-as adatokra vonatkozó) determinációs együtthatója 85,8-ról 84,0%-ra csökken. Ez a csökkenés is alátámasztja, hogy a két távolság közül az első használata a megfelelőbb. Ugyanakkor ennél lényegesebb változás nem várható: az egymással szomszédos megyéknél még jelentősebb lehet az eltérés, de a távolabbi megyéknél egyre kisebbek a különbségek. A kétféle távolság közötti korreláció értéke 0,936.

A két magyarázó változós formát csak 2013-ra megmutatva, a következő eredményt kaptuk:

$$\hat{y}_{ij} = 10^{-7,912} \times \frac{(P_i \times P_j)^{1,132}}{d_{ij}^{1,304}}$$

A távolsághoz tartozó kitevő értéke nem változott. A regresszióhoz tartozó determinációs együttható értéke 85,6%, lényegében szintén megegyezik a 85,8%-kal. Ez a minimális csökkenés indokolt a változók számának a csökkenésével.

Az egy magyarázó változós formával 2013-ban a következő eredményt kaptuk:

$$\frac{\hat{y}_{ij}}{P_i \times P_j} = 10^{-5,827} \times \frac{1}{d_{ij}^{1,606}}$$

Ez jelentősen nagyobb távolsághatást mutat, 1%-kal nagyobb távolság mellett 1,606%-kal csökkent a migráció várt nagysága. Mindemellett a determinációs együttható is jelentős mértékben csökkent, most mindössze 48,5% lett. Ez csak kicsit nagyobb, mint a népességszámmal nem osztott, eredeti adatokkal végzett regresszió 43%-a. Önmagában ez az egy eredmény azt mutatja, hogy az egy magyarázó változós forma nem annyira jól teljesít, mint a két és a három magyarázó változós. Mivel nem találtunk a különböző formák kitevőjének összehasonlíthatóságára vonatkozó korábbi vizsgálatot, a teljes adatmátrixból számos különböző módon hagyunk el területegységeket, és határoztuk meg mindhárom formával a csökkentett adatbázisokra az eredményeket. A három és a két magyarázó változós közötti különbségek elenyészőnek bizonyultak, az egy magyarázó változós viszont rendre lényegesen különbözött a másik kettőtől.

A távolságfüggés összehasonlítása korábbi vizsgálatokkal

Bár a gravitációs modell kalibrálása az adatok szűkössége miatt soha nem számított rutinszerű kutatásnak, az 1960-as évektől kezdődően mégis számos példájával találkozhatunk. Ezek a felhasznált adatok körét tekintve több olyan nagyobb csoportra oszthatók, amelyek mindegyike különböző területekre és térfelosztásokra vonatkozó, de valamennyire mégis összehasonlítható eredményt ad:

1. Egyes országokon belüli migráció.
2. Ingázás, munkába járás.
3. Bevásárlási célú utazások.
4. Turisztikai, szabadidős célú utazások.
5. A lakóhely és a bűnelkövetés helyszíne közötti különbség.
6. A légi közlekedés forgalma.
7. Nemzetközi, országok közötti kereskedelem, valamint országokon belüli áruszállítás.

A fentiekén kívül számos egyéb jelenséget is vizsgáltak, ezek azonban eltérő jellegük miatt egymással nem vagy csak korlátozottan összehasonlíthatók. Részletesebben a népességáramlásra vonatkozó eredményekkel foglalkozunk, amelyek közül néhány szerepel az 5. táblázatban. Illés Iván (1975) számítása kivételével valamennyi vizsgálat a népességszámon és a távolságon kívül bevont különböző demográfiai, gazdasági, földrajzi jellemzőket is. Ilyen jellemzők például a lakosság korszerkezetére, etnikai összetételére, iskolázottságára, gazdasági aktivitására vonatkozó mutatók, továbbá a

migráns lakosság aránya, az átlagjövedelem, a lakásárak, a gazdasági növekedés, a gazdaság ágazati összetétele, a mezőgazdasági vállalkozások átlagos nagysága, a helyi kormányzati kiadások, a városi lakosság aránya, a nagyvárosok léte, a népsűrűség, a klimatikus viszonyokat leíró változók, a területegységek szomszédsága, közbeeső lehetőségek, valamint bármi más, amit az adott kontextusban fontosnak tartanak az elemzés végzői. A távolság azonban a tapasztalatok szerint a változók közül mindig a legjelentősebb mértékben járul hozzá a migráció magyarázatához (Greenwood–Sweetland 1972). Az eredmények többsége azért származik az 1960-as évekből, mert az ilyen típusú vizsgálatok akkoriban voltak divatosak.

Az 5. táblázatban felsorolt eredményekhez több kiegészítést is érdemes fűzni. Illés Iván 1969-es adatokon alapuló számítása egy magyarázó változós formájú volt, így saját eredményeink közül az 1,606-os kitevőjű egy magyarázó változóval vethető össze, és nem a három magyarázó változós, 1,30 és 1,40 közötti értékekkel. Mindenesetre ez az összehasonlítás is a távolsághatás csökkenését mutatja 1969 és 2013 között Magyarországon, a megyék közötti migráció esetében. Indiában a legnagyobb a távolsághatás. Ez részben az ország többnyelvűségéből fakad. A vizsgálat 16 területegysége közül 12 esetében az elsődleges nyelv különbözik az összes többi területegységétől. Nyelvi hasonlóság az egymással szomszédos területegységek között figyelhető csak meg (Greenwood 1971).

A népességszámon és a távolságon kívül más magyarázó változók bevonásának módszertanát itt nem részletezzük. Az eredmények értelmezésével kapcsolatos fontos kérdés, hogy vajon mennyire összehasonlíthatók egymással a különböző magyarázó változókat tartalmazó modellek. Ez a bevont változók körétől és a multikollinearitásra gyakorolt hatásától függ. Utóbbival ritkán foglalkoznak az egyes tanulmányok. A különböző földrajzi dummy változóknak lehet a legnagyobb befolyása a távolsághoz tartozó kitevőre, ugyanakkor a távolságmátrix és a nem térbeli változók között nem lehetséges nagymértékű korreláció. Saját modellünket a kiinduló és a végpont megye átlagbérével kiegészítve például a kitevő abszolút értéke 1,304-ről 1,288-ra módosul, vagyis minimálisan. A kiinduló megye átlagbéréhez tartozó kitevő $-0,718$, a végpont megyéhez $1,043$ tartozik. Vagyis ha a kiinduló megye átlagbére 1%-kal nagyobb, akkor a migráció mértéke $0,718\%$ -kal kisebb lesz (*ceteris paribus*), ha a végpont megye átlagbére 1%-al nagyobb, akkor az áramlás mértéke $1,043\%$ -al nagyobb (*ceteris paribus*).

5. táblázat

A távolsághatás kitevője országos migrációs vizsgálatoknál

Exponent of distance effect at the national migration examinations

Ország és térfelosztás	Kitevő (az egynél nagyobb kitevő egynél több modellt jelent)	A népességszámon és a távolságon kívül egyéb magyarázó változó szerepeltetése a modellben	Irodalom	A felhasznált adatok éve
Magyarország, megye	2,30	nem	Illés I. 1975	1969
India, 16 állam	2,15; 2,01	igen	Greenwood 1971	1961
Ghána, 7 régió	1,43	igen	Beals et al. 1967	1960
Venezuela, 20 állam	1,36; 1,14	igen	Levy-Wadycski 1972	1961
Kína, 31 tartomány	1,20	igen	Li et al. 2014	1985–1990
Egyiptom, 25 régió	1,06	igen	Greenwood 1969b	1960
Egyesült Államok, 48 állam	0,9	igen	Greenwood 1969a	1969, 1972
Német Szövetségi Köztársaság, 11 tartomány	0,80; 0,81; 0,84; 0,86	igen	Vedder et al. 1971	1967
Brazília, 19 állam	0,79	igen	Sahora 1968	1950
Egyesült Államok, 48 állam	0,72; 0,83; 0,84	igen	Lin-Xie 1998	1990
Egyesült Államok, 50 nagyvárosi régió	0,702	igen	Greenwood–Sweetland 1972	1955–1960
Jamaica, 13 régió	0,66	igen	Adams 1969	1960
Egyesült Államok, 100 ezernél népesebb nagyvárosok	0,62	igen	Stouffer 1962	1935–1940
Indonézia, 26 tartomány	0,22; 0,62	igen	Lottum–Marks 2012	1930
Egyesült Államok, 250 ezernél népesebb nagyvárosok	0,42	igen	Galle–Taeuber 1966	1955–1960
Kína, 31 tartomány	0,21	igen	Li et al. 2014	1990–1995
Kína, 31 tartomány	0,10	igen	Li et al. 2014	1995–2000

6. táblázat

A távolsághatás kitevője néhány nem migrációs vizsgálatnál
 Exponent of distance effect at some non-migration investigations

Vizsgálat	Kitevő	Irodalom
Kiskereskedelmi forgalom	0 – 12,5	Reilly 1929
Telefonhívások néhány egyesült államokbeli város körzetében	2,3 – 3,3	Carroll 195
Telefonhívások, Malajzia	0,94	Leinbach 1973

A 6. táblázat a számos további jelenség áramlására vonatkozó sokféle eredményből hármat tartalmaz. Ezek közül elmélettörténeti fontossága és az elmélettörténet által többnyire torzított módon bemutatott volta miatt kiemelendők Reilly (1929) a kiskereskedelmi forgalom nagyságára vonatkozó kalibrációi. Reilly összesen 255 kalibrációt végzett el különböző településhármasok között (két kereskedelmi központnak egy harmadik településre gyakorolt vonzerejére vonatkozóan), vagyis annyiból nem illeszkedik a korábbi példák közé, hogy nem egyetlen kitevőt számított egyetlen területegységre. Számításainál a kitevők a 0 és a 12,5 közötti tartományba kerültek. Ezek szám-tani átlaga 3,49, a móduszt tartalmazó osztályköz a 1,5–2,5-ig terjedő intervallum volt. Reilly ennek az osztályköznek a közepét választotta „tipikus” kitevőnek, ami pont megegyezik a fizikai analógia kitevőjével, így még vonzóbbnak bizonyult az 1940-es években népszerűvé váló társadalomfizika képviselői számára. Reilly későbbi értelme-zői, hivatkozói pedig rendre megfelelkeznek a kitevő nagymértékű ingadozásáról, amikor Reilly kiskereskedelmi vonzástörvényével és annak kettes kitevőjével foglal-koznak, ahelyett, hogy felhívnák a figyelmet a tapasztalati adatokon alapuló újrakalib-rálás fontosságára. Reilly eredményeinek nagymértékű szóródása pontosan azt mu-tatja, hogy számszerű értelemben nem létezik kiskereskedelmi vonzástörvény.

Összegzés

Az 5. és a 6. táblázatok legfontosabb tanulsága a kitevőkre vonatkozó eredmények nagymértékű különbözősége. Ez jól indokolható a jelenségek eltérő térérzékenysége-vel, az eltérő időszakokkal, területi alapegységekkel és térfelosztásokkal, migrációs de-finíciókkal, és általában azzal, hogy az emberek aktív cselekedetei koncepciónálisan teljesen eltérnek a fizikai világ alkotóelemeinek passzív viselkedésétől. A gravitációs modell érénye abban rejlik, hogy képes bemutatni a tömeges emberi viselkedésre vo-natkozóan kimutatható szabályosságokat, a szabályosság mértékét a determinációs együttható nagysága révén számszerűsíti, valamint objektív módon – időben, térben és jelenségek között – összehasonlíthatóvá teszi a távolsághatás jelentőségét. Ehhez azonban tapasztalati áramlási adatokra és azokon alapuló állandó újrakalibrálásra van szükség.

IRODALOM

- ADAMS, N. A. (1969): Internal Migration in Jamaica: An Economic Analysis *Social and Economic Studies* 18 (2): 137–151.
- BAJMÓCY P.–KISS, J. P. (1999): Megyék, régiók és központjaik – modellek tükrében *Tér és Társadalom* 13 (1-2): 31–51.
- BEALS, E. E.–LEVY, M. B.–MOSES, L. N. (1967): Rationality and Migration in Ghana *Review of Economics and Statistics* 49 (4): 480–486.
- BELLA, B. (2011): Balatonvilágos dilemmája: Veszprém vagy Somogy megye *Comitatus* 21 (2): 50–59.
- BELUSZKY, P. (1974): *Nyíregyháza vonzáskörzete* Földrajzi Tanulmányok. 13. Akadémiai Kiadó, Budapest.
- CARROLL, J. D. (1955): Spatial Interaction and the Urban-Metropolitan Regional Description *Papers and Proceedings of the Regional Science Association* 1 (1): 59–73.
- GALLE, O. R.–TAEUBER, K. E. (1966): Metropolitan Migration and Intervening Opportunities *American Sociological Review* 31 (1): 5–13.
- GREENWOOD, M. J. (1969a): An Analysis of the Determinants of Geographic Labor Mobility in the United States *Review of Economics and Statistics* 51 (2): 189–194.
- GREENWOOD, M. J. (1969b): The Determinants of Labor Migration in Egypt *Journal of Regional Science* 9 (2): 283–290.
- GREENWOOD, M. J. (1971): An Analysis of the Determinants of Internal Labor Mobility in India *The Annals of Regional Science* 5 (1): 137–151.
- GREENWOOD, M. J.–SWEETLAND, D. (1972): The Determinants of Migration between Standard Metropolitan Statistical Areas *Demography* 9 (4): 665–681.
- HARDI, T.–SZÖRÉNYINÉ KUKORELLI, I. (2014): The Extension of Complex Gravity Zones and Manpower Catchment Areas in North Transdanubia In: CSIZMADIA, Z.–DUSEK, T. (szerk.): *The Győr Automotive District* pp. 39–62., Széchenyi István Egyetem, Universitas-Győr Nonprofit Ltd, Győr.
- ILLÉS, I. (1975): *Regionális gazdaságtan* Tankönyvkiadó, Budapest.
- KINCSES, Á.–TÓTH, G. (2012): Gravitációs modell alkalmazása a térszerkezet vizsgálatára *Terrületi Statisztika* 52 (5): 479–491.
- KINCSES, Á.–TÓTH, G. (2014): The Application of Gravity Model in the Investigation of Spatial Structure *Acta Polytechnica Hungarica* 11 (2): 5–19.
- KISS, J. P.–BAJMÓCY, P. (2001): Városi funkciójú központok és elméleti vonzáskörzeteik az Alföldön *Tér és Társadalom* 15 (1): 65–89.
- LACKÓ, L. (1978): Települések vonzásterületének meghatározása egymásrahatási modell segítségével *Földrajzi Értesítő* 27 (1): 31–43.
- LEINBACH, T. R. (1973): Distance, Information Flows, and Modernization: Some Observations from West Malaysia *Professional Geographer* 25 (1): 7–11.
- LEVY, M. B.–WADYCKI, W. J. (1972): Lifetime versus One-Year Migration in Venezuela *Journal of Regional Science* 12 (3): 407–415.
- LI, Y.–LIU, H.–TANG, Q. (2014): Interprovincial Migration in China, 1985-2005: a Gravity Modelling Approach *Chinese Journal of Population Resources and Environment* 12 (2): 146–156.

- LIN, G.–XIE, Y. (1998): The Loglinear Modeling of Interstate Migration: Some Additional Considerations *American Sociological Review* 63 (6): 900–907.
- LOTTUM, J. VAN–MARKS, D. (2012): The Determinants of Internal Migration in a Developing Country: Quantitative Evidence for Indonesia, 1930–2000. *Applied Economics* 44 (34): 4485–4494.
- NAGY, G. (1996): A gravitációs modell alkalmazási lehetőségei a településen belüli mozgások tanulmányozására *Tér és Társadalom* 10 (2–3): 149–156.
- NAGY, G. (2011): A gravitációs modell felhasználásának lehetőségei a várostérségek lehatárolásában *Területi Statisztika* 51 (6): 657–673.
- NEMES NAGY, J.–PIROS, GY. (1984): Térbeli népesség- és anyagáramlások vizsgálata gravitációs modellel In: SIKOS, T. T. (szerk.): *Matematikai és statisztikai módszerek alkalmazási lehetőségei a területi kutatásokban* pp. 171–175., Akadémiai Kiadó, Budapest.
- NEUMANNÉ VIRÁG, I. (2014): *Az integráció hatása az EU tagországok külkereskedelmére – vizsgálata a gravitációs modellel* Doktori értekezés, Pannon Egyetem, Veszprém
- PAPP, A. (1981): Debrecen vonzáskörzete *Alföldi Tanulmányok* 5: 177–204.
- PÉNZES, J. (2005): Városi vonzásközpontok vizsgálata az Észak-alföldi régióban In: SÜLI-ZAKAR, I. (szerk.): „Tájak – Régiók – Települések...” Tiszteletadás a 75 éves Enyedi György akadémikus előtt pp. 160–165., Didakt Kft., Debrecen.
- REILLY, W. J. (1929): *Methods for the Study of Retail Relationships* University of Texas, Austin
- SAHOTA, G. S. (1968): An Economic Analysis of International Migration in Brazil *Journal of Political Economy* 76 (2): 218–245.
- STOUFFER, S. A. (1962): *Social Research to Test Ideas* Free Press of Glencoe, New York.
- SZALKAI, G. (2012): A járások kialakításának módszertani megalapozása *Területi Statisztika* 52 (3): 215–229.
- SZIVA, M. (1987): *Kísérlet Győr vonzáskörzetének meghatározására gravitációs modellel* In: MTA RKK kutatási eredményei. Térszerkezeti vizsgálatok az Észak-Dunántúlon pp. 58–74., MTA RKK, Pécs.
- TAGAI, G. (2011): *Térkapcsolati modellek a regionális kutatásokban* Doktori értekezés, ELTE, Budapest.
- TÓTH, G.–KINCSES, Á. (2007): Közúti elérhetőségi vizsgálatok Európában *Statisztikai Szemle* 85 (5): 431–463.
- TÓTH, G. (2005): A potenciálmodell alkalmazásának lehetőségei az autópálya-nyomvonalak területfejlesztési szempontú vizsgálatában *Gazdaság és Statisztika* 56 (3): 3–18.
- VEDDER, R. K.–GALLAWAY, L. E.–CHAPIN, G. L. (1971): The Determinants of Internal Migration in West Germany, 1967 *Weltwirtschaftliches Archiv* 106: 309–317.