УДК 004.6, 004.9, 681.51

О. С. Кушнір *, С. В. Рихлюк *, Я. П. Кісь
Національний університет "Львівська політехніка",
кафедра інформаційних систем та мереж;
Львівський національний університет імені Івана Франка,
*кафедра оптоелектроніки та інформаційних технологій

СИСТЕМА НАСЕЛЕНИХ ПУНКТІВ УКРАЇНИ ТА ЇЇ ЧАСОВА ЕВОЛЮЦІЯ

Вивчено часову еволюцію статистичних розподілів розмірів населених пунктів України і параметрів домінування основного міста за 1897–2013 роки. В широкому діапазоні загальноукраїнські та регіональні залежності ранг–населення підкорюються оберненому степеневому закону Ціпфа з близьким до одиничного значенням степеня. Діапазон найменших розмірів населення, найімовірніше, описується логнормальною функцію. Степеневий характер рангової залежності для українських міст засвідчує риси самоорганізації, ієрархії та скейлінгу, притаманні складним системам. Встановлено часову стабільність форми залежності ранг–населення та показано, що статистичний розподіл віку українських міст має помірно асиметричний характер, а кореляція між населенням і віком міст відсутня. Доведено, що рангові залежності для різних функціональних типів населених пунктів України не описуються законом Ціпфа.

Ключові слова: складні системи, закон Ціпфа, степеневий розподіл, логнормальний розподіл, розподіл розмірів міст, часова еволюція.

Time evolution of the statistical distribution of Ukrainian settlement sizes and the parameters of dominance of the main city are studied for 1897-2013. Both the national and regional rank-size dependences are governed by the inverse-power Zipf's law in a wide range, with the exponent close to unity, while the region of the smallest population sizes is likely described by the lognormal function. A power-law character of the rank dependence for the Ukrainian cities certifies the features of self-organization, scaling and hierarchy inherent to the complex systems. It is found that the rank-size dependence is stable in time, the statistical distribution of the city ages has a moderately asymmetric character, and the city age and its population are not correlated. The rank dependences for different functional types of Ukrainian settlements are not described by the Zipf's law.

Key words: complex systems, Zipf's law, power-law distribution, lognormal distribution, city size distribution, time evolution.

Вступ

На відміну від найпростіших природних систем, яким часто притаманні ґауссовий або близькі до нього статистичні розподіли з "легкими хвостами", складні системи та мережі часто виявляють степеневі розподіли, які слугують своєрідним статистичним "відбитком пальців" цих систем [1]. Зокрема, соціально-економічна поведінка – це яскрава ілюстрація складності, з багатьма типами (іноді нелінійних) взаємодій між елементами системи. Одним із прикладів складних систем є сукупності міст (або, дещо загальніше, населених пунктів), які в разі обмеження аналізу певною країною чи регіоном формують відносно замкнуті системи, в межах яких компоненти мають істотні економічні, транспортні, соціальні та багато інших функціональних зв'язків. Дослідники почали цікавитися статистикою розподілу розмірів міст доволі давно – ще починаючи з робіт Ф. Ауербаха

(1913), А. Лотки (1925) і Ґ. Ціпфа (1941) (див., наприклад, праці [2, 3]). Проте й зараз продовжуються напружені дискусії стосовно різних проблем. Це загальні теоретичні передумови закону Ціпфа для залежності ранг–населення і зростання міст за законом пропорційності Р. Ґібрата [4–9], конкретні чинники, що зумовлюють ієрархію міст (економічні, географічні, політичні, екологічні [10] або суто статистичні [6]), переваги і недоліки логнормального та степеневого розподілів в описі статистики поселень [5, 7, 11, 12], належні методичні процедури для оцінювання цієї статистики [1, 13], а також альтернативні дефініції "міст" (адміністративні, функціональні тощо) [3, 10, 14].

Дилема про застосовність логнормального чи оберненого степеневого статистичних розподілів до опису населення міст украй принципова (див. [3, 15]). За центральною граничною теоремою статистики, загальним наслідком мультиплікативного стохастичного процесу та закону Ґібрата (незалежності розмірів міста x і швидкості його відсоткового зростання) повинні бути "рандомна прогулянка" $\log x$ і його нормальний розподіл (тобто, логнормальний для x). З іншого боку, еволюцію складних систем і мереж часто описують механізмом "багатий стає багатшим" і теоретичною моделлю "переважного приєднання" Г. Саймона (див. [16, 17]), які повинні породжувати степеневий розподіл В. Парето із "важким хвостом" [4], самоорганізовану складність системи міст та їхнього населення, скейлінг у цій системі та її своєрідну масштабну інваріантість або самоподібність. Нарешті, незважаючи на деяку зовнішню схожість (див. [18]), логнормальний і степеневий розподіли насправді докорінно різні: усі моменти першого з них скінченні, а другий має лише обмежену кількість скінченних перших моментів. Інакше, степеневий закон Парето для статистичного розподілу національного багатства, стану фінансових ринків, розмірів фірм та інших параметрів економічної діяльності [19] "нечесний" – поняття на зразок "середнього багатства людини" або "населення середнього міста" часто не мають змісту (див. [14]). Іноді це стосується навіть відхилення від середнього, яке теж може не існувати в строгому математичному розумінні. Ось чому вивчення статистики населених пунктів актуальне не лише для прогнозування процесів урбанізації, але й в ширшому фундаментальному плані.

Окрім наведених вище міркувань, наше дослідження стимулювали ще три такі аспекти.

1. Як влучно зазначив автор праці [20], переважна більшість досліджень за предметом стосується США. Є хіба одиничні приклади порівняння обмеженої кількості міст для багатьох країн – це дещо застарілі дані [21] (1980 р.) і нещодавні дані [22] (2005 р.). Докладні ж результати для більшості інших країни практично відсутні. Зокрема, за наявною в авторів інформацією, питання розподілу населення міст такої крупної європейської країни як Україна (населення 42,7 млн. станом на 2016 рік) зачіпали лише в порівняльному аналізі [22], у праці [23] і в єдиному вітчизняному дослідженні [24].

2. У літературі переважно вивчали населення міст, хоча відомо, що найбільші відхилення від степеневого закону може виявляти саме слабше вивчений "хвіст" рангової залежності, який містить найменші за розмірами поселення [5, 7, 11, 12, 15]. У зв'язку з цим ми розглядатимемо і міста, і малі населені пункти України.

3. Хоча зазвичай дослідники постулюють для системи міст властивість ергодичності (тобто закономірності стохастичного процесу зростання в часі деякого міста еквівалентні закономірностям для вибірки багатьох міст у деякий момент часу), все ж тісно пов'язані питання часової еволюції розподілу розмірів міст і взаємозв'язку їхніх віку та розмірів сьогодні висвітлено доволі скупо, а відповідні результати суперечливі. Так, автори [2, 25] знайшли статистично значиму кореляцію віку і населення міст США, а висновки праці [26] по суті протилежні.

У зв'язку з цим мета цієї роботи – масштабний статистичний аналіз складної системи населених пунктів України у рамках комп'ютерних методів, з'ясування співвідношення логнормальної та степеневої функцій в описі залежності ранг–розмір населених пунктів, з'ясування кореляції між їхнім віком і населенням, а також аналіз закономірностей і масштабів скейлінгу на хвості рангової залежності.

Емпіричні дані та методи досліджень

Для вивчення динаміки розвитку населених пунктів розглянемо результати переписів населення, які на території України за останні 80 років відбувалися в 1937, 1939, 1959, 1970, 1979, 1989 і 2001 роках (див. дані [27, 28]). З перших двох украй близьких за часом переписів ми обмежилися більш недавніми даними за 1939 рік. Ми залучили до розгляду дані перепису в Російській імперії (див. [29]), які, щоправда, не охоплюють усіх українських територій, а також свіжіші статистичні дані за 2011–2013 роки [27]. Дані переважно стосуються лише міських поселень (міст і с. м. т.). Винятком є 2001 рік, дані за який найповніші і містять дані про населення всіх міст, селищ міського типу (с. м. т.), селищ і сіл.

Дані, представлені в друкованому форматі, було оцифровано. Для нагромадження і зберігання результатів було створено базу даних, а для їхнього статистичного опрацювання – програмний пакет в середовищі Visual Studio. Зокрема, коефіцієнт кореляції Пірсона R будь-яких аналізованих статистичних параметрів u і v ми розраховували безпосередньо за його визначенням, без застосування стандартних математичних пакетів. Параметр R якнайкраще описує лінійний зв’язок параметрів u–v, а насправді цей зв’язок може мати істотно складніший (степеневий, експоненційний тощо) характер. Тому додатково ми вивчали кореляцію в координатах u–log$v$, log$u$–v і log$u$–log$v$. Ці самі підходи сприяли згладжуванню можливих викидів функції $v(u)$ (див. [6]).

Найпростішим представленням ієрархії населених пунктів є залежність $x(r)$ їхнього населення x залежно від рангу r населених пунктів. Значення $r = 1$ відповідає населеному пунктові із найбільшим розміром населення, $r = 2$ – другому за населенням і т. д. У табл. 1 наведено загальні дані, що характеризують розподіл населення України в різні роки – повну кількість населених пунктів, яка дорівнює найвищому рангові r_{max}, а також розмір (тобто населення) x_{min} найменшого з поселень, охопленого нашими даними.

Таблиця 1

Деякі статистичні дані для міських поселень України за різні роки (кількість населених пунктів r_{max} і розміри x_{min} найменшого з них) і параметри апроксимації степеневою функцією залежностей ранг–населення (коефіцієнт Ціпфа α, модуль коефіцієнта кореляції R, поріг апроксимації $s = x_{0min}$)

Рік	Параметр											
	s = 5 тис.		s = 10 тис.		s = 20 тис.		s = 50 тис.		s = 100 тис.		r_{max}	x_{min}
	α	R	α	R	α	R	α	R	α	R		
1897 [1]	0,84	0,993 [2]	**0,80** [3]	0,996	0,85	0,993	0,87	0,984	0,93 [4]	0,989 [4]	138	1215
1939	**0,96**	0,998	0,99	0,997	0,98	0,995	0,96	0,986	0,90	0,961	549	1028
1959	**0,91**	0,998	0,94	0,997	0,96	0,995	0,92	0,984	0,83	0,970	1103	190
1970	**1,00**	0,998	**1,00**	0,998	0,99	0,996	0,92	0,993	0,86	0,987	1241	198
1979	**1,03**	0,999	1,03	0,998	1,03	0,995	0,96	0,992	0,87	0,989	1285	140
1989	**1,05**	0,999	1,04	0,998	1,03	0,995	0,96	0,991	0,88	0,985	1329	75
2001 [5]	**1,02**	0,998 [6]	1,04	0,997	1,03	0,994	0,97	0,989	0,86	0,981	29790 [5]	1
2011	**1,04**	0,998	1,05	0,997	1,03	0,994	0,98	0,987	0,86	0,979	1342	43
2012	**1,04**	0,998	1,07	0,996	1,03	0,994	0,98	0,986	0,86	0,979	1342	27
2013	**1,04**	0,998	1,05	0,997	1,03	0,994	0,98	0,986	0,86	0,979	1343	3

Примітки:

[1] Дані стосуються територій сучасної України, які на момент перепису входили до Російської імперії

[2] Залежність $x(r)$ в подвійному логарифмічному масштабі виявляє незначні нелінійності

[3] Жирний шрифт виділяє дані апроксимації, яка характеризується найвищим коефіцієнтом R

[4] Апроксимація ґрунтується на даних лише для чотирьох міст

[5] Дані за 2001 рік описують усі населені пункти України, а не лише міста і с. м. т., як для решти років

[6] Для значення порогу s = 2 тис. одержуємо ще вище значення R = 0,999 і α = 0,99 (див. текст)

Гіпотеза про масштабно інваріантну обернену степеневу (ціпфівську) рангову залежність відповідає формулі

$$x(r) = Cr^{-\alpha}, \qquad (1)$$

де $\alpha \sim 1$ – це показник степеня Ціпфа, C – стала нормування, значення якої часто ототожнюють з розміром найбільшого міста ($C = x_1 = x_{\max}$ – див. також обговорення [13, 30]). На практиці випадкову змінну, пов'язану з розмірами x населених пунктів, можна вважати неперервною. Виконання формули (1) найпростіше перевірити за фактом лінійності залежності ранг–населення в подвійному логарифмічному масштабі

$$\log x = \log C - a \log r\,. \tag{2}$$

Якість лінійної апроксимації в координатах log$x$–log$r$ визначається коефіцієнтом кореляції R і нормованою сумою квадратів відхилень.

Дещо складнішим методом є перехід від рангової залежності $x(r)$ до еквівалентного представлення кумулятивної функції розподілу $P(x)$ (або залежності Парето). З точністю до масштабного коефіцієнта остання дорівнює функції, оберненій до рангової:

$$r(x) = Ax^{-k}\,, \tag{3}$$

де k – коефіцієнт Парето. Потім переходять до густини розподілу $p(x) = Bx^{-b}$, де $\beta \sim 2$ – це константа. Тоді для неперервного розподілу $p(x)$ наявні аналітичні формули для знаходження степеня β, одержані за стандартним методом максимальної подібності [30]. Відомі також загальні співвідношення між наведеними вище коефіцієнтами: $k = 1/\alpha$, $\beta = 1 + 1/\alpha$.

Альтернативою у статистиці поселень є гіпотеза про логнормальний розподіл імовірності їхніх розмірів:

$$p(x) = 1/(\sqrt{2p}\, s x) e^{-(\log x - m)^2/2s^2}\,, \tag{4}$$

де μ і σ – це середнє значення та середньоквадратичне відхилення змінної x, відповідно. Тоді рангова залежність описуватиметься оберненою кумулятивною функцією нормального розподілу $\Phi^{-1}(\log x)$ (див. [9]). Характерною ознакою логнормального характеру рангової залежності є нелінійність кривої $x(r)$ в координатах log$x$–log$r$, тобто непостійність кутового нахилу $\alpha(r)$, із поступовим збільшенням $\alpha(r)$ зі зростанням r (див. формули (2) і (4)).

З літератури відомо (див. розділ 1), що для залежностей ранг–населення типовою є ситуація, коли ділянка найбільших населень x і найменших r ("голова") описується степеневою функцією, а ділянка найменших x ("хвіст"), а іноді й центральна ділянка ("тіло") – логнормальною функцією. Мінімальне значення $x_{0\min}$, при якому ще не помітні нелінійності залежності log$x$(log$r$), а коефіцієнт R залишається достатньо великим, тоді відповідатиме "порогові" $s = x_{0\min}$, після якого (при $x \geq x_{0\min}$) опис формулою (1) коректний.

Крім загальнонаціональних даних, ми опрацювали також регіональні дані для ієрархії населених пунктів України за 2001 рік, почерпнуті з джерел [27, 28]. Крім того, визначалися параметри "домінування" p_1 і p_2 головного міста, визначені авторами праці [21]:

$$p_1 = x_1 \left(\sum_{i=1}^{5} x_i \right)^{-1},\ p_2 = x_1 \left(\sum_{i=1}^{50} x_i \right)^{-1}, \tag{5}$$

де $x_1 = x_{\max}$ – це населення відповідно першого за чисельністю міста, а $\sum_{i=1}^{5} x_i$ і $\sum_{i=1}^{50} x_i$ – сумарне населення відповідно перших п'яти і п'ятдесяти населених пунктів країни. Оскільки природною гіпотезою було би пов'язати "крутішу" залежність log$x$(log$r$), тобто вищий степінь α, із більшою (або меншою) роллю перших (або останніх) за рангом населених пунктів, ми розраховували коефіцієнти кореляції Пірсона між параметрами α, p_1 і p_2.

Ще одним предметом нашого інтересу було вивчення даних за віком міст України станом на поточний рік. Зазначимо, що міста, для яких була наявна інформація лише про століття, у якому засновано місто, але не відповідний рік, було опущено. У результаті із загальної кількості понад 450 міст оброблялися результати для 411 міст України. У знаходженні масової функції розподілу віку українських міст і розрахунку середнього віку та середньоквадратичного відхилення використано прийом логарифмічного бінування.

Результати та їхнє обговорення
Еволюція ієрархії населених пунктів України в національному масштабі

Рангові залежності для населених пунктів України міського типу за 1897, 1939, 1959, 1970, 1979, 1989, 2001, 2011–2013 роки наведено на рис. 1, а. Там же наведено повніші дані 2001 року для всіх населених пунктів. На рис. 1,б показано лінійну апроксимацію залежності log$x$(log$r$) за даними 2001 року лише для міського населення. Нарешті, на рис. 1, в проілюстровано часову еволюцію ціпфівського параметра α, розрахованого для порогів населення s = 5, 10, 20, 50, 100 тис., разом із еволюцією параметрів домінування головного міста p_1 і p_2. Тут обрання різних порогів (тобто різних діапазонів $x_{0\min}\div x_{\max}$ лінійної апроксимації функції log$x$(log$r$)) дає змогу вивчити стабільність і якість параметрів апроксимації, а також встановити можливу залежність цих параметрів від порогу. Пороги s, які відповідають степеневій залежності, параметри α і R лінійної апроксимації log$x$(log$r$) за різні роки, а також загальну кількість аналізованих населених пунктів $r_{\max}$ і розміри $x_{\min}$ найменшого з них відображено в табл. 1.

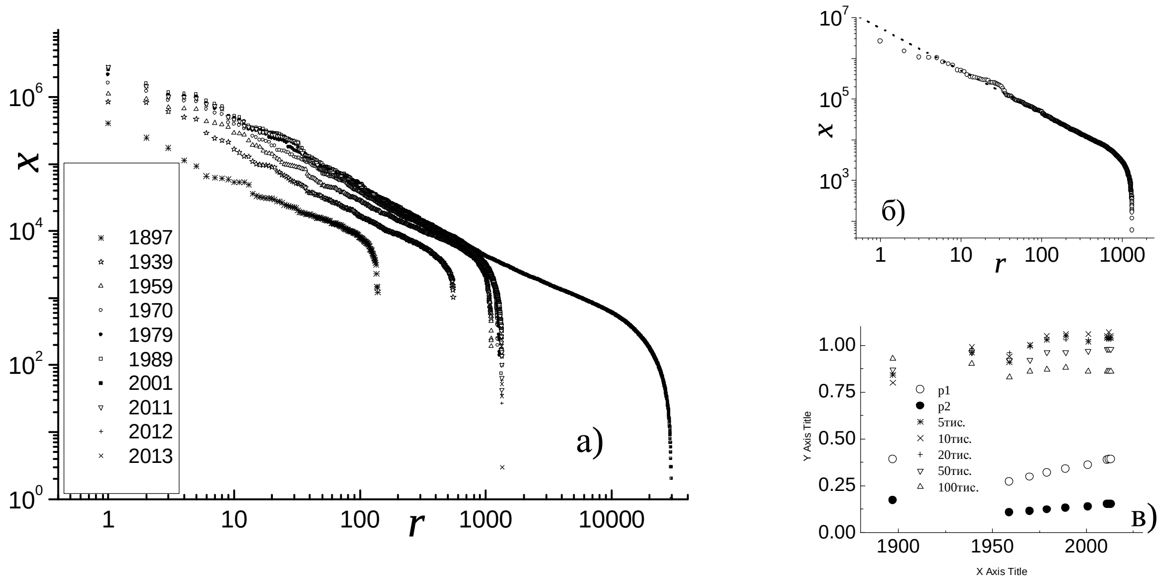


Рис. 1. Залежності ранг–населення для міських поселень України (дані за 2001 рік відповідають усім населеним пунктам) (а); ця сама залежність для міських поселень України за 2001 рік та її лінійна апроксимація (штрихована лінія) при порозі населення 10 тис. (б); часова еволюція ціпфівського параметра α (для порогів s = 5, 10, 20, 50, 100 тис.) і параметрів домінування p_1 і p_2 (в)

Дані рис. 1,а спонукають до висновку, що значення порогу населення впливає на величину параметра Ціпфа α. Окрім даних за 1897 рік, величина α дещо зменшується зі зростанням порогу (див. табл. 1). Водночас зменшується і коефіцієнт R, який визначає якість апроксимації (див. подальше обговорення). Тому найбільшу довіру викликають степені α, виділені в табл. 1 жирним шрифтом, які переважно стосуються порогу s = 5 тис. Ці степені представлені також на рис. 1, в, разом із альтернативними величинами α, які відповідають більшим порогам (10, 20, 50 і 100 тис.).

Висновок про вплив порогу на показник степеневої залежності Ціпфа узгоджується з літературою (див. наприклад, [22]). Проте конкретні значення s, одержані для різних країн у різних роботах, доволі істотно відрізняються. Так, автор праці [22] переважно обирав поріг s = 10 тис., окрім великих країн із більшою кількістю населення. У дослідженнях міст США [1, 12] знайдено пороги переходу від степеневої до логнормальної рангової залежності відповідно s = 60 і 37 тис.; у праці [20] типові пороги для США становили 50, 60 або й 100 тис.; у [15] для Англії та Уельсу знайдено відповідно s = 35 і 50 тис., а в [11, 21] для більшості країн поріг становив 100 тис.

Для України оптимальний у плані мінімізації параметра R поріг ~ 5 тис., проте пороги 10 і 20 тис. дають близькі результати для степеня α. Скінченна точність апроксимації рангової залежності степеневою функцією і деяка залежність $a(s)$ вказують, що подавати дані для a (або відповідні дані k для оберненої залежності Парето $r(x)$) з надто високою точністю (наприклад, до 3-

х або 4-х знаків після коми, як у працях [21, 22]) не коректно. Нарешті, для найдокладніших даних за 2001 рік оптимальний поріг ще нижчий ($s = 2$ тис. – див. примітку в табл. 1), точність апроксимації вкрай висока (0,999), а нелінійності $\log x(\log r)$ практично непомітні навіть для аномально низьких порогів $s = 0{,}5$ і 1 тис. ($\alpha = 0{,}90$, $R = 0{,}998$ і $\alpha = 0{,}93$, $R = 0{,}997$, відповідно).

Ми припускаємо, що поріг $s = 2$ тис. виявився би оптимальним і для інших досліджених років, якби ми оперували для них повними даними для всіх населених пунктів, а не лише даними для міст і с. м. т. За нашою інформацією, таких широких діапазонів степеневої ділянки залежності ранг–населення в літературі досі не було згадано для жодної з країн. Рангова залежність для всіх населених пунктів України за найповнішими даними 2001 року лінійна в подвійному логарифмічному масштабі по суті на чотирьох декадах. Отже, ми доходимо висновку про аномально низький поріг виконання степеневого закону розподілу населення України. Оскільки важко пристати на думку про абсолютну унікальність залежності $x(r)$ для України, автори передбачають потребу в уважному перегляді закономірностей ієрархії населених пунктів інших країн з урахуванням статистики всіх поселень. Нарешті, степінь a для України дуже близький до класичного ціпфівського значення $a = 1$, статистичну надійність якого в універсальному описі всіх країн багато дослідників відкидають (див. [21, 22]).

Зазвичай степеневу ділянку рангової залежності $x(r)$ для великих x вважають важливою не через відповідну кількість міст, а через значний відсоток населення, що мешкає в них. Так, кількість міст США з розмірами понад 100 тис., ієрархію яких описує степенева функція, становить лише 1 % від загальної кількості всіх населених пунктів; щоправда, в них проживає 63 % населення країни [12] (див. також схожі оцінки в праці [1]). Незважаючи на слабшу урбанізацію України, поріг для неї істотно нижчий, а тому для 2001 року ($s = 2$ тис.) виявляється, що до степеневої ділянки залежності $x(r)$ слід віднести не лише понад 74 % населення, але й понад 8 % усіх населених пунктів. Отже, для України ця ділянка набуває не лише принципового, але й практичного значення.

Окрім повних даних за 2001 рік, відхилення від лінійності для всіх решти рангових залежностей у подвійному логарифмічному масштабі стають помітнішими при порогах, менших за 5 тис. Водночас, зі зростанням порогу понад 5 тис. і звуженням діапазону апроксимації коефіцієнт кореляції R дещо парадоксально також зменшується (див. табл. 1). На нашу думку, це пов’язано з домінуванням у процедурі апроксимації основної маси експериментальних точок у “тілі” залежності $x(r)$ і на її “хвості”, а також двома специфічними рисами ієрархії населених пунктів України. По-перше, це дещо занижене, порівняно з передбаченнями формули Ціпфа, населення перших чотирьох міст (Києва, Харкова, Одеси та Дніпропетровська), майже незалежно від обраного порогу. Наприклад, населення Києва за цими передбаченнями мало б сягати $x_1 \approx 5{,}3$ млн., а не 2,6 млн., як насправді (див. рис. 1б). На нашу думку, “недонаселеність” перших міст України легко врахувати, уводячи поправку Б. Мандельброта [30] до закону Ціпфа (1):

$$x(r) = C(r + r_0)^{-a}\,, \tag{6}$$

де r_0 – стала, не надто відмінна від одиниці. Формула (6) асимптотично еквівалентна до (1) при $r >> 1$, але дещо зменшує теоретично передбачуване населення перших міст ($r \sim 1 \div 10$).

Інша специфіка залежності $x(r)$ для України – це немонотонність залежності $x(r)$ у діапазоні $r = 15 \div 35$ і “перенаселеність” (знову ж таки, порівняно з прогнозом за законом Ціпфа) міст із цими рангами, що найліпше видно з рис. 1, б. Отже, зі звуженням діапазону апроксимації (підвищенням порогу) зростає внесок точок, пов’язаних з цими містами, а тому якість апроксимації дещо понижується, а кутовий нахил α зменшується.

Цікаво порівняти наші дані з іншими результатами для України. У дослідженні [23] для України було вказано лише параметри нелінійності рангової залежності $\log x(\log r)$ (див. нижче). У праці [22] було вивчено 103 міста України (як стверджує автор, за даними 1998 року) і одержано $k \approx 1{,}02$, тобто $\alpha \approx 0{,}98$. Це значення близьке до наших даних 0,96 за 1989 рік і 0,97 за 2001 рік, якщо припустити, що поріг тут становив $s \approx 50$ тис. (див. табл. 1). Більше того, пряма перевірка 103-х за рангами населених пунктів у найближчі до 1998 роки (1989 і 2001) показує, що їхнє населення становить 45 і 51 тис., відповідно. Отже, поріг у дослідженні [22] для України справді дорівнював 50 тис., а дані [22] досконало узгоджуються з нашими. Нарешті, у праці [22] наведено значення параметра $A \approx 7{,}0$ млн., який фігурує в формулі (3). Проте, твердження [20] про рівність $A = x_1$ неправильне, крім спеціального випадку $\alpha = 1$ (див. [15, 22]). Загалом же маємо $A = x_1^{1/a}$, звідки навіть при такому близькому до одиниці значенні α, як 0,98 [22] для Києва одержуємо помітно менше населення ($x_1 \approx 6{,}7$ млн.). Наші ж менші α дадуть ще нижчі та коректніші оцінки.

У роботі [24], де джерел і часової прив'язки вихідних даних не вказано, було опрацьовано дані для 404 населених пунктів і одержано $x_1 \approx 9{,}1$ млн., $\alpha \approx 1{,}15$ і $R \approx 0{,}953$. Величини обох параметрів x_1 і α викликають сумніви, а коефіцієнт лінійної кореляції істотно нижчий порівняно з нашими даними. Проте позитивом дослідження [24] був розрахунок залишків апроксимаційної процедури, які загалом виявилися помітними та становили до 150 % для найбільших міст. На додаток, очевидно неґауссовий характер залежності залишків від рангу [24] підтверджує або присутність похибок невідомого походження в даних [24], або відомі обмеження опису рангових залежностей населення міст степеневою функцією. Останнє узгоджується зі знайденим нами фактом залежності коефіцієнта a від вибору порогу s, а тому з деякою випуклістю функції $x(r)$. У зв'язку з цим окремі дослідники [8, 23, 22] пропонують перейти до загальніших функцій на зразок

$$\log x = \log A' + a' \log r + b' (\log r)^2 + \ldots, \tag{7}$$

де A', a' і b' – деякі константи. Ми не використовували такого підходу, оскільки за виразом (7) для функції $x(r)$ не стоїть жодна теоретична гіпотеза або конкретна модель ієрархії населених пунктів.

Параметри домінування першого міста p_1 і p_2 для України за останні десятиліття близькі до даних [21], наприклад, для Італії, Канади, Німеччини, Польщі та Швейцарії. Величини p_1 для України за 1980 рік (0,36÷0,39 – див. рис. 1в) помітно менші, ніж відповідні дані, наприклад, для США (0,45), Радянського Союзу (0,47), Австрії (0,70) та Угорщини (0,76); величини p_2 (0,14÷0,15 – див. рис. 1, *в*) теж менші, ніж для США (0,19), Угорщини (0,45) і Австрії (0,50). Отже, Україні важко приписати виражене явище домінування столиці та значну "нерівність" населення різних міст. Це узгоджується і з "недонаселеністю" столиці та кількох інших найбільших міст (див. вище).

Часто урбаністи говорять про примат столиць і великий коефіцієнт Ціпфа α (малий коефіцієнт Парето k), притаманні системі міст тоталітарних країн (див. [31], стор. 32–37 і, частково, [22]). Проте аналіз і порівняння даних [21, 22] для різних країн (зокрема, Радянського Союзу, Росії, США тощо), а також аналіз наших даних для різних історичних епох розвитку України не підтверджує такого зв'язку. Так, коефіцієнти α (0,96 і 0,91) і параметри p_1 (0,26 і 0,27) за 1939 і 1959 роки виявляються навіть дещо меншими, ніж за недавні роки, а параметр домінування $p_1 = 0{,}39$ часів царської Росії (1897 рік) – істотно вищим, аніж для сталінських часів. Цікаво також порівняти дані $k = 1{,}28$ (тобто $\alpha = 0{,}78$) [21] для всього Радянського Союзу за 1980 рік із нашими даними для України фактично на цей самий момент часу (1979 рік). Навіть найменше зі значень коефіцієнта α, одержане нами для порогу 100 тис. ($\alpha = 0{,}87$ – див. табл. 1), істотно відмінне від 0,78. Це доводить помітну регіональну неоднорідність системи міст Радянського Союзу. У цьому плані заслуговує на дослідження припущення про своєрідну нестійкість цього по суті наддержавного утворення щодо системного підходу.

Заслуговує на інтерес питання, яке орієнтовне місце системи міст України серед інших країн світу. На 1980 рік автори [21] одержали "середньосвітове" значення $\bar{k} \approx 1{,}14$ (тобто $\bar{a} \approx 0{,}88$), користуючись порогом степеневої залежності $s = 100$ тис. Наші значення α в діапазоні 0,86÷0,88 для України останніх десятиліть і цього самого порогу практично збігаються зі згаданим середнім. Залежно від методу статистичного оцінювання, у масштабному порівнянні різних країн [22] переважно за даними від 2000 року знайдено $\bar{k} \approx 1{,}11$ або $\bar{k} \approx 1{,}17$, тобто $\bar{a} \approx 0{,}90$ або $\bar{a} \approx 0{,}86$. Усереднюючи, одержимо значення $\bar{a} \approx 0{,}88$. Оскільки поріг тут становив 50 тис. (див. вище), наше відповідне значення для України за 2001 рік (0,97 – див. табл. 1) приблизно на 10 % більше за $\bar{a}$. Отже, загалом систему населених пунктів України можна назвати типовою.

Ми визначили кореляцію $R \approx 0{,}92$ між двома параметрами домінування p_1 і p_2, яка добре узгоджується з даними [21] ($R \approx 0{,}90$). Загальні міркування підказують, що параметри домінування повинні також корелювати з крутизною рангової залежності в подвійному логарифмічному масштабі, тобто з коефіцієнтом Ціпфа α. Це підтверджує і виявлена авторами [21] антикореляція коефіцієнта Парето k, для якого відомий зв'язок $k = 1/\alpha$, із параметрами p_1 і p_2. Ми перевірили ці результати на прикладі наших даних α, знайдених із порогом $s = 10$ тис. для різних років. Одержані дані ($R_{a-p_1} = 0{,}09$ і $R_{a-p_2} = -0{,}26$) не підтвердили кореляції, проте причиною виявилася перша часова точка за 1897 рік. Відповідним даним бракує системності, оскільки вони стосуються не всієї України (див. [23] і розділ 1). Вилучення єдиної точки за 1897 рік різко змінило результати: кореляція R між α для порогу 10 тис. і p_1 або p_2 становить 0,89 або 0,88, відповідно. Звісно, наші статистичні дані кількісно вбогіші, бо стосуються не багатьох різних країн, як у праці [21], а лише десяти різних років для єдиної країни. Однак вони таки підтверджують кореляцію крутизни степеневої рангової залежності та параметрів домінування першого міста в системі міст країни [21].

Часова еволюція ціпфівського параметра для України відсутня або зводиться до кількісно неістотної тенденції до зростання (див. табл. 1 і рис. 1, в). Середнє значення параметра α в 1897–2013 роках становить 0,99, а стандартне відхилення не перевищує 7 %. Після 1970 року відповідні цифри становлять 1,03 і 1,7 %. Інакше кажучи, за майже 120 років рангова функція $x(r)$ у більшості аспектів залишилася незмінною. Отже, система населених пунктів України демонструє значну і довготривалу стабільність, яку не порушили украй серйозні, як за мірками багатьох країн, політичні та соціально-економічні катаклізми – дві світові війни, тоталітарна урбанізація, сталінський геноцид, а також перехід до ринкової економіки, довготривала криза та негативні тенденції приросту населення. Ці факти схиляють до думки, що походження скейлінгу $x(r)$ пов'язано не з економікою, політикою, екологією чи географією, а швидше із деякими найзагальнішими закономірностями статистики складних систем [32] (див. також обговорення в підрозділі 3.4).

Порівняно з часовою залежністю параметра Ціпфа, еволюція параметрів домінування виразніша кількісно (див. рис. 1, в): за той самий аналізований період маємо зростання p_1 і p_2 на 30–50 %. Мабуть, ця тенденція збережеться і в найближчі десятиліття.

3.2. Регіональні закономірності ієрархії населених пунктів

Окрім рангових залежностей загальнонаціонального масштабу, ми вивчали також регіональні закономірності ієрархії населених пунктів України. На рис. 2 показано приклади залежностей ранг–населення для чотирьох регіонів України за даними 2001 року. У табл. 2 наведено коефіцієнт Ціпфа α, модуль коефіцієнта кореляції R і поріг s апроксимації залежностей ранг–населення для поселень усіх типів у шести регіонах України за той самий рік. Рис. 2 проілюструє лінійну апроксимацію $\log x(\log r)$ лише для тих порогів населення, що забезпечують найвищі коефіцієнти кореляції для кожного регіону.

Як видно з табл. 2, величини a для різних регіонів України змінюються в неочікувано широкому діапазоні (0,73–1,53), що контрастує із даними для всієї країни за більш ніж столітній період часу. Цей факт засвідчує помітну неоднорідність ієрархії населених пунктів у регіональному масштабі. Це важливе явище потребує докладного обговорення, яке представлено в підрозділі 3.3. Для регіонів помітне майже монотонне зростання коефіцієнта Ціпфа a зі зростанням обраного порогу s, тобто зі звуженням діапазону апроксимації $x_{0\min} \div x_{\max}$. Ця тенденція набагато виразніша і за своїм характером протилежна до знайденої в підрозділі 3.1 закономірності для загальнонаціонального розподілу населення. Можливо, її причина полягає в тому, що для досліджених регіонів спостерігаємо "перенаселеність" кількох перших за рангом міст (див. апроксимаційні прямі на рис. 2). З іншого боку, наявність чіткої закономірності в поведінці $a(s)$ породжує припущення про те, що регіональні рангові залежності можуть насправді виявитися близькими до логнормальних (див. обговорення в підрозділі 3.3). Точніше інтерпретувати ці залежності слід за допомогою строгих методів нелінійної апроксимації.

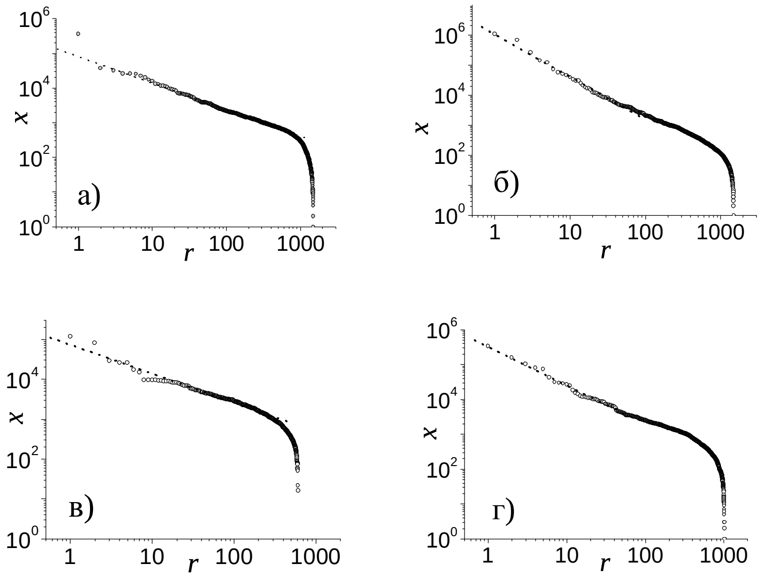


Рис. 2. Приклади залежностей ранг–населення для різних регіонів України за даними 2001 року: Вінницька обл. (а), Дніпропетровська обл. (б), Закарпатська обл. (а) і АР Крим (г). Штриховані лінії – лінійна апроксимація в подвійному логарифмічному масштабі для порогів населення s = 0,5 тис. (а), 4 тис. (б), 1 тис. (в) і 4 тис. (г)

Таблиця 2

Параметри апроксимації степеневою функцією залежностей ранг–населення для регіонів України за 2001 рік (коефіцієнт Ціпфа α, модуль коефіцієнта кореляції R і поріг апроксимації s)

Регіон	s = 0,5 тис.		s = 1 тис.		s = 4 тис.		s = 10 тис.	
	α	R	α	R	α	R	α	R
Вінницька обл.	**0,76**	0,997	0,79	0,992	0,90	0,970	0,98	0,926
Волинська обл.	0,73	0,980	0,88	0,977	1,24	0,987	**1,42**	0,985
Дніпропетровська обл.	1,14	0,994	1,22	0,994	**1,42**	0,996	**1,53**	0,996
Житомирська обл.	0,79	0,983	0,94	0,984	1,20	0,989	**1,38**	0,992
Закарпатська обл.	0,82	0,980	**0,73**	0,990	0,75	0,971	1,16	0,970
АР Крим	0,92	0,993	0,91	0,991	**1,10**	0,994	1,19	0,993

Примітка. Жирний шрифт виділяє дані апроксимації, яка характеризується найвищим коефіцієнтом R

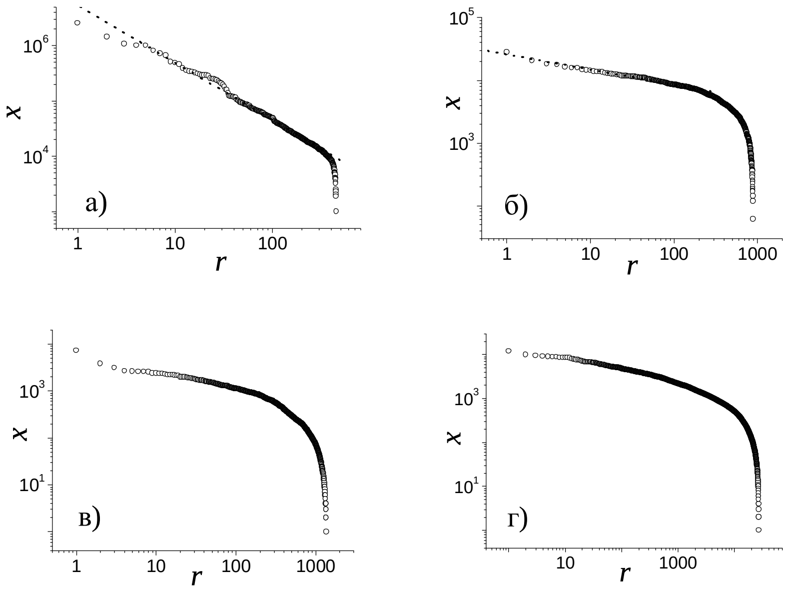
Нарешті, на відміну від глобальної по країні тенденції до пониження параметра R зі зростанням порогу s, на регіональному рівні спостерігаємо дещо хаотичний характер поведінки $R(s)$. Зокрема, найбільші значення R для різних регіонів досягаємо при різних порогах (див. табл. 2). Проте, як і в національному масштабі, про виконання інтуїтивно очікуваної тенденції до зростання R зі звуженням діапазону апроксимації залежностей $\log x(\log r)$ теж не йдеться. Мабуть, і тут загальною причиною є ефект скінченних розмірів вибірки – адже з лінійним зростанням порогу кількість населених пунктів зменшується експоненційно.

Рангові залежності для різних типів населених пунктів

Перейдемо до аналізу рангових залежностей для "функціонально" різнотипних населених пунктів – міст, с. м. т., селищ і сіл. Як видно з рис. 3, для міст і порогу s = 15 тис. лінійна апроксимація залежності $\log x(\log r)$ дає α = 1,04, R = 0,996, що узгоджується з даними табл. 1 для 2001 року. Проте навіть для с. м. т., які функціонально найближчі до міст, результати несподівані: хоча якість апроксимації формально достатня (R = 0,994), а її діапазон (s = 7 тис.) начебто належить степеневій ділянці рангової залежності на рис. 1, *а*, *б*, ми одержуємо α = 0,24. Для селищ і сіл спроби лінійної апроксимації приводять до схожих результатів: відповідно, α = 0,35, R = 0,987 і α = 0,27, R = 0,987. Їх навряд чи можна інтерпретувати як наявність степеневих ділянок на рангових залежностях, хоча би через надто низькі ціпфівські степені α.

Отже, спроби описати степеневою функцією залежності ранг–населення для поселень окремих, функціонально різних типів марні, крім випадку міст. На відміну від результатів для національного та регіонального рівнів, обидві залежності $a(s)$ і $R(s)$ монотонні, тобто розширення діапазону лінійної апроксимації $\log x(\log r)$ понижує якість останньої та систематично збільшує розрахований коефіцієнт Ціпфа. Ми припускаємо, що насправді дані для с. м. т., селищ і сіл описуються логнормальною функцією. Загальні причини такої докорінно іншої поведінки можна збагнути на підставі системного підходу: на відміну від поселень деякого регіону країни, а тим більше на відміну від усієї сукупності населених пунктів деякої країни, такі складові цих систем, як с. м. т., селища і села принципово не формують самостійних замкнутих систем.

Рис. 3. Залежності ранг–населення для різних типів населених пунктів України за даними 2001 року, які охоплюють 451 місто (а); 888 с. м. т. (б), 1359 селищ (в) і 27092 сіл (г). Штриховані лінії – лінійна апроксимація в подвійному логарифмічному масштабі для міст (s = 15 тис.) і с. м. т. (s = 7 тис.)



Висвітливши здобуті нами результати для України загалом, її окремих регіонів і функціонально різних типів поселень, ми можемо порівняти їх з відомими теоретичними моделями для ієрархії населених пунктів (див., наприклад, [2, 3, 9, 11, 12]). Населені пункти формуються внаслідок стохастичних процесів природного відтворення та міграції населення, а також реструктурування (поява, зникнення або злиття) поселень [9]. Хоча суто мультиплікативний характер цих процесів приводить до логнормального розподілу ймовірності $p(x)$ розмірів населених пунктів, для появи степеневого "хвоста" $p(x)$ достатньо хоча би незначної за масштабами присутності ("домішки") адитивного стохастичного процесу [2, 17]. Тоді великі та малі населені пункти матимуть різні механізми зростання: у розвитку великих міст переважно домінуватимуть флуктуації приросту населення, а для малих поселень визначальними будуть флуктуації параметрів міграції [9]. За цих умов одержуємо степеневу "голову" рангової залежності $x(r)$ та її логнормальний "хвіст", що пояснює основні результати цієї роботи. Проте додаткового кількісного пояснення вимагає факт аномально широкого діапазону розмірів населених пунктів x, у якому рангова залежність $x(r)$ для України має степеневий характер.

На думку авторів [15], на регіональному рівні ми маємо переважно логнормальну за характером ієрархію поселень, а глибинною причиною появи степеневої ділянки на кривій $x(r)$ для всіх населених пунктів деякої країни слугують локальні відмінності масштабів урбанізації, тобто деякий розподіл регіональних параметрів. Інакше, ціпфівська залежність є результатом "суміші" локальних логнормальних розподілів, а також логнормальних розподілів для різних "функціональних категорій центрів" [15]. На жаль, поняття цих категорій у праці [15] конкретизовано недостатньо. Ми припускаємо, що може йтися про поселення на зразок міст, селищ, сіл тощо. У такому разі наші результати, зокрема факти значних регіональних відмінностей рангових залежностей і специфічних кривих $x(r)$ для різних функціональних типів поселень, повністю підтверджують припущення авторів [15].

У працях [3, 11] було введено універсальну функцію для опису рангової залежності для всіх розмірів x – це подвійний парето-логнормальний розподіл, “тіло” якого логнормальне, а обидва краї мають степеневий характер. Важливо, що механізм появи згаданого розподілу такий самий – це відмінності локальних розподілів [3, 11]. Як ілюстрацію автор праці [11] наводить дані для кількох американських штатів і провінцій в Іспанії, які засвідчують наявність степеневих ділянок залежності $x(r)$ при малих і великих x. Проте опис кривих $x(r)$ для України подвійною парето-логнормальною функцією наштовхується на труднощі: ми не виявили лінійної ділянки на жодній із досліджених кривих $\log x(\log r)$ в області “хвоста” $r_{\max}$, навіть для найменших за розмірами поселень $x_{\min} \sim 1 \div 100$ (див. рис. 1, 2 і 3).

Кореляція населення і віку українських міст

Для з’ясування розподілу віку міст України було розраховано функцію густини ймовірності $p(t)$ віку міст України. Через недостатній обсяг статистичної вибірки неперервної змінної t і відповідні шуми було використано логарифмічне бінування функції. Оптимальна кількість інтервалів (бінів), яка мінімізувала ці шуми, становила 10. Результати наведено на рис. 4, а. Середній вік $\bar{t}$ українських міст дорівнює $\bar{t} = 443$ роки, а відповідне середньоквадратичне відхилення – $\Delta t = 325$ років. Цікаво також, що сумарний вік усіх проаналізованих міст понад 180 тис. років. Розподіл $p(t)$ має помірно асиметричний, псевдопуассонівський характер, із коефіцієнтом асиметрії (відхилення від розподілу на зразок нормального) $w = 0{,}60$. Цікаво, що для такої молодої країни, як США автори [2, 25] одержали, з урахуванням поправки на час опублікування цих досліджень, $\bar{t} \approx 131$ рік і $\Delta t \approx 50$ років.

Наступним завданням було вивчення можливої залежності населення міст від їхнього віку. Знайдені нами коефіцієнти кореляції Пірсона для функціональної залежності $x(t)$ складають $R = 0{,}08$, $-0{,}03$, $0{,}05$ і $-0{,}02$ відповідно для представлення в масштабах t–x, t–$\log x$, $\log t$–x і $\log t$–$\log x$. Малі абсолютні величини кореляції та залежність її знаку від масштабу по осях абсцис і ординат засвідчують, що населення і вік українських міст насправді не пов’язані між собою – статистична кореляція фактично відсутня.

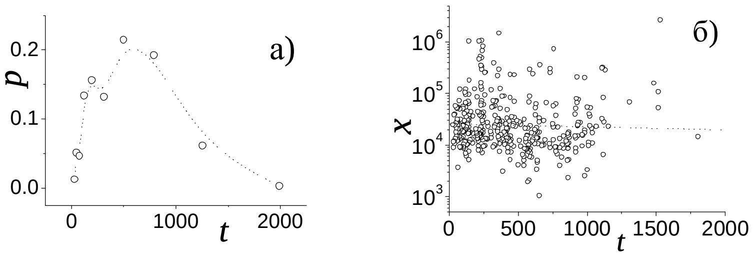


Рис. 4. Функція густини імовірності p(t) для віку міст України (а); кореляція між їхнім населенням x і віком t у масштабі t–logx (б)

На перший погляд, наші результати суперечать даним робіт [2, 25], де було виявлено статистично значиму кореляцію віку і населення міст США і, залежно від кількості врахованих міст, одержано коефіцієнти кореляції від 0,57 до 0,65. Проте ми схильні припускати, що знайдена в [2, 25] кореляція випадкових змінних x і t, мабуть, пояснюється порівняно малим (~ 300 років), а також, можливо, порівняно “однорідним” (за відсутністю радикальних подій тощо) дослідженим часовим проміжком. На часових шкалах, що відповідають більшим історичним періодам, така кореляція цілком може зникнути. Так, автор [26] розрізняє зовнішню динаміку або макродинаміку – динаміку розподілу або ієрархії об’єктів (тут – міст), які входять до системи, – а також внутрішню мікродинаміку, за якою ці об’єкти змінюють розміри і ранг. На значному статистичному матеріалі в роботі [26] доведено, що мікродинаміка істотна на іноді дуже різних часових і просторових шкалах: міста можуть швидко або повільно розвиватися та занепадати, як розвиваються та занепадають навіть цілі імперії або й цивілізації, проте ця мікродинаміка типово прихована за зовнішньою стабільністю макророзподілу об’єктів у системі. Відповідно, ці дві складові добре узгоджуються з нашими результатами для еволюції системи міст України та зв’язку віку і розмірів міст. З одного боку, внаслідок нестабільної мікродинаміки старші міста зовсім необов’язково найбільші за населенням; з іншого, оскільки макророзподіл у системі майже не залежить від часу, рангова залежність для всієї системи міст країни залишається по суті незмінною.

Висновки

Аналізуючи емпіричні дані для кластеризації населення України та ієрархію розмірів цих кластерів, виявлено відмінне виконання закону Ціпфа на ділянках "голови" та "тіла" залежності ранг–населення, а також підтверджено можливість теоретичного опису її "хвоста" логнормальною функцію. Відповідно, механізми зростання малих і великих населених пунктів повинні бути різними. Степенева рангова залежність для порівняно великих за розмірами населених пунктів засвідчує, що сукупність українських міст повинна виявляти основні риси самоорганізації, ієрархії та масштабного скейлінгу, що притаманні складним системам.

Дані дослідження функцій x(r) для національного та регіональних масштабів засвідчують, що коефіцієнт Ціпфа α дещо залежить від статистичної вибірки населених пунктів і порогу населення s. Оптимальний поріг для опису міських поселень становить $s \approx 5$ тис., для всіх населених пунктів – 2 тис., а для регіональних рангових залежностей – залежить від регіону. Степенева ділянка залежності x(r) для України аномально широка і простягається на чотири декади. Степінь Ціпфа α для України близька до класичного значення $\alpha = 1$.

Розраховано параметри домінування основного міста в ієрархії міст України та виявлено фактичну відсутність тенденції до домінування столиці. Описано часову еволюцію степеневого параметра Ціпфа α для українських поселень і параметрів домінування Києва протягом більш ніж 100 років. За останні майже 40 років залежність ранг–населення виявляє значну стабільність форми; незалежно від вибору порогу відносне збільшення показника степені α не перевищують $1 \div 4$ %. Виявлено статистично значущу позитивну кореляцію між параметрами домінування p_1 і p_2 першого міста, а також між обома цими параметрами і степенем Ціпфа α.

Доведено задовільний (хоча й дещо гірший, аніж у національному масштабі) опис залежностей ранг–населення законом Ціпфа для різних регіонів України. Уперше показано, що рангові залежності x(r) для функціонально різних підгруп населених пунктів України на зразок с. м. т., селищ і сіл із переважно малими розмірами не описуються законом Ціпфа. Ці емпіричні факти, які можна збагнути в межах системного аналізу, слід належно врахувати в кількісній теорії, яка описує граничний перехід від логнормального до оберненого степеневого розподілу розмірів населених пунктів за умови їхнього зростання. Наші дані для найменших поселень не виявили степеневої ділянки на жодній із досліджених залежностей x(r), а тому вони не підтверджують припущення про т. зв. подвійний парето-логнормальний розподіл. Показано, що статистичний розподіл віку українських міст має помірно асиметричний псевдопуассонівський характер, а кореляція між населенням і віком міст України фактично відсутня.

Предметом подальшої роботи автори вбачають точнішу кількісну інтерпретацію рангових залежностей розмірів населених пунктів України за методами нелінійної апроксимації, зокрема встановлення статистичної коректності нульової гіпотези про її ціпфівський характер на ділянках "голови" і "тіла" шляхом перевірки нормального розподілу залишків.

1. Malevergne Y., Pisarenko V., Sornette D. // Phys. Rev. E. – 2011. – Vol. 83. – 036111 (11 p.). 2. Giesen K., Suedekum J. // SERC Discussion Paper 122. – UK Spatial Economics Research Centre (SERC). – 2012. – 25 p. 3. Berry B. J. L., Okulicz-Kozaryn A. // Cities. – 2012. – Vol. 29. – P. S17–S23. 4. Zanette D. H., Manrubia S. C. // Phys. Rev. Lett. – 1997. – Vol. 79, No 3. – P. 523–526. 5. Eeckhout J. // American Econom. Rev. – 2004. – Vol. 94, No.5. – P. 1429–1451. 6. Li Gan, Dong Li, Shunfeng Song. // Economics Lett. – 2006. – Vol. 92. – P. 256–262. 7. Levy M. // American Econom. Rev. – 2009. – Vol. 99, No.4. – P. 1672–1675. 8. Benguigui L., Blumenfeld-Lieberthal E. // J. Geogr. Syst. – 2011. – Vol. 13. – P. 87–100. 9. Kaldasch J. // ISRN Economics. – 2014. – Article ID 498125. – 6 p. 10. Decker E. H., Kerkhoff A. J., Moses M. E. // PLOS ONE. – 2007. – Is. 9. – e934 (6 p.). 11. Reed W. J. // J. Region. Sci. – 2002. – Vol. 42, No 1. – P. 1–17. 12. Ioannides Y., Skouras S. // J. Urban Econom. – 2013. – Vol. 73. – P. 18–29. 13. Urzúa C. M. – Working Paper EGAP-2010-04, Tecnológico de Monterrey, Campus Ciudad *de México. – 2010. – 4 p. 14. Bin Jianga, Tao Jia. // Int. J. Geograph. Inform. Sci. – 2011. – Vol. 25, No. 8. – P. 1269–1281. 15. Parr J. B., Suzuki K. // Urban Studies. – 1973. Vol. 10. – P. 335–352. 16. Barabási A.-L. // Nature. – 2005. – Vol. 435. – P. 207–211. 17. Zanette D. H. // Adv. Complex Syst. – 2008. – Vol. 4/11. – P. 1–16. 18. Perline R. // Statist. Sci. – 2005. – Vol. 20, No. 1. – P. 68–88. 19. Axtell R. L. // Science. – 2001. – Vol. 293, No 5536. – P. 1818–1820. 20. Veneri P. – In: OECD Regional Development Working Papers, 2013/27. – OECD Publishing, 2013. – 16 p. – Режим доступу: http://dx.doi.org/10.1787/5k3tt100wf7j-en 21. Rosen K. T., Resnick M. // J. Urban Econom. – 1980. – Vol. 8. – P. 165–186. 22. Soo K. T. // Regional Sci. Urban Econom. – 2005. – Vol. 35. – P. 239–263. 23. Benguigui L., Blumenfeld-Lieberthal E. // Computers, Environment and Urban Systems. – 2007. – Vol. 31. – P. 648–666. 24. Ібатуллін Ш. І. // Містобудування та терит. планув. – 2006. – Вип. 24. – С. 72–80. 25. Giesen K., Suedekum J. // European Econom. Rev. – 2014. – Vol. 71. – 193–208. 26. Batty M. // Nature. – 2006. – Vol. 444. – P. 592–596. 27. Кількість та територіальне розміщення населення України за даними Всеукраїнського перепису населення 2001 р. – Ред. О. Г. Осауленка, відп. за вип. Н. С. Власенко, Л. М. Стельмах. – К : Держ. комітет статист. України, 2003. – 218 с. 28. Статистика населення України. – Режим доступу: http://database.ukrcensus.gov.ua 29. Демоскоп Weakly. – Ин-т демографии Нац. исслед. ун-та “Высшая школа экономики”. – Режим доступу: http://demoscope.ru/weekly/ssp/rus_lan_97_uezd.php?reg=426 30. Newman M. E. J. // Contemp. Phys. – 2005. – Vol. 46. – P. 323–351. 31. Becker C., Mendelsohn S. J., Benderskaya K. Russian urbanization in the Soviet and post-Soviet eras. – Urbanization and Emerging Population Issues, Working Paper 9. – International Institute for Environment and Development, United Nations Population Fund. – 2012. – 134 p. 32. Adamic L. // Nature. – 2011. – Vol. 474. – P. 154–165.*