

НОВИ ВЪЗМОЖНОСТИ ЗА ИНОВАЦИИ В СТАТИСТИЧЕСКАТА ТЕОРИЯ И ПРАКТИКА - ГЕО КОДИРАНЕ НА СОЦИАЛНИТЕ ИЗСЛЕДВАНИЯ

*Валерия Ангелова**



Въведение

През последните години интересът и вискателността на потребителите към статистически данни на ниво по-ниско от националното трайно нараства. Данни на ниво NUTS 2 и 3, общини и населени места, наричани Local Administrative Units (LAU), квартали, преброителни участъци, сгради, адреси, басейнови дирекции, заливни тераси на реки, грид клетки, както и специализирани функционални географски райони и териториални типологии, основаващи се на грида на населението, са само част от примерите. Сами по себе си грид клетките, използвани в статистическата практика, варират по площ в зависимост от целите на конкретното изследване и природата на наблюдаваното явление. Така например грид клетки с площ 10 квадратни километра са подходящи за целите на статистиката на емисиите на вредни вещества в атмосферата, но са твърде големи за нуждите на демографската и социалната статистика, тъй като концентрацията и социално-икономическият профил на населението може силно да

* Магистър по регионално развитие и управление, държавен експерт в отдел „Регионална статистика и индикатори за мониторинг“, Национален статистически институт; e-mail: VAngelova@nsi.bg.

варира на сравнително малки по площ територии. Въз основа на дългогодишни проучвания и обсъждания на европейско ниво грид клетките с площ от 1 квадратен километър са приети като еталон за целите на демографската и социалната статистика, въпреки че някои държави като Нидерландия и Швейцария разполагат с по-детайлен грид на населението (страна от 100 метра). Независимо от площта на клетката основното предимство е, че грид клетките са устойчиви във времето и са независими от промените в административните и административно-териториалните граници на страните.

Както в българската, така и в европейската практика най-широко използвана за представяне и разпространение на данни на ниво по-ниско от национално е Класификацията на териториалните единици за статистически цели (от френски *La nomenclature des unités territoriales statistiques* (NUTS)). Тя е създадена въз основа на административните структури на държавите - членки на Европейския съюз (ЕС), и се основава на принципа за йерархичност и праг за брой население. Границите на тези единици често са исторически унаследени и не отразяват териториалната структура на социалната и икономическата реалност в България и Европа. Според Регламента за NUTS границите на регионите биха могли да се актуализират на всеки три години¹. По правило всяко изменение и допълнение на Регламента за NUTS води до утвърждаване на регламент за изпълнение относно установяването на еднакви условия за предаване на динамичните редове за новото регионално разпределение. Условието, зададено от регламента, се базира на консултации със заинтересованите страни и производителите на данни от националните статистически институти. Според последния документ, отразяващ промените в NUTS класификацията, който е в сила от 13 ноември 2020 г., държавите членки имат задължението да предоставят на Евростат преизчислени данни по новите граници на NUTS 3 регионите от следните тематични области: демография - население, раждания на живо дете и умиряния (назад до 1990 г., като обаче подаването на данни за годините преди 2012 г. не е задължително), регионални сметки - сметки по отрасли и по региони (до 2000 г.) и пазар на труда - заетост и безработица (до 2016 г.). Това означава, че разполагаме с различна дължина на динамичните редове по показатели, свързани с две ключови направления, използвани за планиране и мониторинг на териториалната кохезия (БВП и състояние на пазара на труда). Въпреки че за преизчисление на данните до ниво 2 по NUTS са предвидени повече тематични области в сравнение с преизчислението по NUTS 3, резултатът и в този случай е прекъсване на

¹ <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02003R1059-20180118&qid=151913675347>

динамичните редове. Например в домейните „туризъм“ и „околна среда“ - съоръжения за преработка на отпадъци - е задължителна само тригодишна ревизия на данните, а ревизията на данните в сферата на здравеопазването и информационното общество е само на доброволен принцип. Допълнително усложнение за националните и европейските институции, отговорни за създаването и имплементирането на политики с териториално измерение, е фактът, че задължението за предоставяне на динамичните редове е до 1 януари 2023 г., т.е. две години след влизането в сила на основния регламент.²

Що се отнася до българската практика още по-комплициран е въпросът относно ефекта от промените на границите на местните административни единици. На първо място няма установени международни правила относно кодирането и прекодирането на местните административни единици. НСИ ежегодно докладва на Евростат актуалния списък на LAU компонентите в състава на NUTS 3 регионите посредством т. нар. LAU листи³. При промяна в границите на LAU кодовете в тези списъци, както и кодовете, поддържани в Единния класификатор на административно-териториалните и териториалните единици (ЕКАТТЕ), остават непроменени⁴. В динамичните редове, публикувани на страницата на НСИ, няма индикация за промяна в границите на местните административни единици. Като пример могат да се посочат годишните данни за механичното движение на населението по общини. Това може да доведе до колебания на данните във времето, труднообясними за потребителите, които не са статистици. Пространственият ефект от промените в границите на общините е онагледен в следващата секция от анализа.

Целта на настоящото изследване е да се анализират практически неограничените възможности на грида на населението за осигуряване на още по-разнообразни субнационални данни с по-дълги динамични редове. Добавена стойност за статистическата теория и практика може да бъде реализирана и по отношение на повишаването на качеството на извадките на социалните изследвания. Анализът съчетава аспекти от административно-териториалното, териториалното и

² <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32020R1703>

³ <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units>

⁴ ЕКАТТЕ се поддържа в рамките на Националния регистър на населените места, който е изграден от Националния статистически институт, съгласувано с Министерството на регионалното развитие и благоустройството и Министерството на земеделието, храните и горите в изпълнение на чл. 37, ал. 2 от Закона за административно-териториалното устройство на Република България. Регистърът е достъпен на адрес <https://www.nsi.bg/nrnm/>

функционалното географско райониране за статистически и аналитични цели, прилагането на териториалните типологии в различни национални и европейски политики, въпроси, свързани с теорията на извадките и възможностите на методите за оценки на данни за малки териториални единици (от английски - Small Area Estimation Methods (SAE)). Поставените въпроси са актуални в дългосрочен план, тъй като очертаните в статията предизвикателства ще излизат на дневен ред при всяко прерайониране на страната и Европа.

1. Относно ефектите от периодичните промени в границите на териториалните единици за статистически цели

В България границите на NUTS регионите не са актуализирани от 2006 г. при последна налична версия на класификацията от 2021 година. В други държави след 2006 г. настъпват съществени административно-териториални промени. Например в Германия през 2013 и 2016 г. регионите от ниво NUTS 3 претърпяват значителни модификации, в Гърция, Франция, Полша, Португалия, Словения и Обединеното кралство - през 2013 г., в Ирландия, Нидерландия, Полша, Финландия и Обединеното кралство - през 2016 г., а в Белгия, Италия, Обединеното кралство и Норвегия - през 2021 година. Други държави като Литва, Унгария и Хърватия инициират промени само на ниво NUTS 2⁵. В неотдавнашни изявления на министри от служебния кабинет (през юни 2021 г.) беше споменато, че е целесъобразно да последваме опита на други европейски страни и да отделим столицата в самостоятелен NUTS регион от ниво 2.

Освен преизчисление на данните всички тези промени изискват прекодиране на регионите, своевременно информиране на потребителите и имплементирането на новите резултати в националните бази данни и в базата данни на Евростат. Всичко това представлява допълнително натоварване за Евростат и производителите на официална статистика на национално равнище. Към това следва да прибавим затруднението, което потребителите изпитват при работа с регионални данни с различната дължина на динамичните редове по домейни.

Следва да споменем и друга подробност, полезна за потребители, които комбинирано използват регионални данни от страницата на НСИ и тази на Евростат, а именно използването на различни кодове за 28-те области. Например по ЕКАТТЕ кодът

⁵ <https://ec.europa.eu/eurostat/web/nuts/background>

на област Благоевград е BLG, а по Регламента за NUTS НСИ докладва данните за областта с код BG413.

На общинско равнище проблемът с прекъснатите динамични редове поради промени в границите е особено осезаем. Според Бояджиев и Веселинова: „Историята на вътрешните граници между териториалните единици на страната показва, че те са нестабилни и постоянно се променят. Най-стабилни се оказват границите между отделните селища. Най-нестабилни са границите между териториалните равнища под среден ранг.“⁶ С цел да илюстрирам пространствените трансформации в границите на общините в резултат на прехвърлянето на населени места от една община в друга избрах примера на общините от Южното Черноморие на страната, тъй като това е територия, където се наблюдават едни от най-динамичните промени в границите през годините. Представените промени обхващат периода от 1995 до 2010 година. Разглежданият период е подбран така, че да включва всички промени в състава на общините от региона при стабилни граници на област Бургас. Последната промяна, свързана със създаването на нова община с център Обзор, не е включена в изследването, тъй като към момента на писане на настоящата статия няма издаден указ на президента на Република България. На фиг. 1 е представен първоначалният състав на общините по Южното Черноморие на България към 31.12.1995 година. С Указ № 383/обн., ДВ, бр. 98/4.11.1996 г. селата Маринка и Твърдица, които дотогава са в границите на община Созопол, преминават към община Бургас (фиг. 2)⁷. Най-мощабните административно-териториални промени произтичат от Указ № 258/обн., ДВ, бр. 57/11.07.1997 г., като село Димчево преминава от община Созопол към община Бургас, град Приморско и село Писменово излизат от състава на община Царево и стават част от новосъздадената със същия указ община Приморско, а селата Ясна поляна и Ново Паничарево преминават от община Созопол в община Приморско (фиг. 3). По-късно, вследствие на Указ № 102/обн., ДВ, 17.04.2001 г. (фиг. 4) град Китен преминава от състава на община Царево към този на община Приморско. С Указ № 127/обн., ДВ, 24.04.2002 г. село Веселие преминава от община Созопол към община Приморско (фиг. 5). Последната активна промяна в

⁶ Бояджиев, В., Н. Веселинова (2017). Границите като проблем за регионалното развитие на България. В: Сб. Юбилейна международна научна конференция БЪЛГАРИЯ НА РЕГИОНИТЕ Перспективи за устойчиво регионално развитие, Пл.; <http://regions.uard.bg> (31.08.2021 г.; 15:00) с. 557.

⁷ На фигури от 2 до 6 имената на общините, които губят населени места за сметка на съседни общини, са обозначени в червено. Авторът изказва специални благодарности на г-н Димитър Попов от НСИ за подготовката и предоставянето на картограмите.

административно-териториалното деление на общините по Южното Черноморие е в резултат на Указ № 121/обн., ДВ, бр. 38/21.05.2010 г., според който село Извор излиза от състава на община Созопол и преминава към община Бургас (фиг. 6).

Фиг. 1. Населени места в състава на общините по Южното Черноморие на България към 31.12.1995 година



Източник: НСИ.

Фиг. 2. Населени места в състава на общините по Южното Черноморие на България към 5.11.1996 година



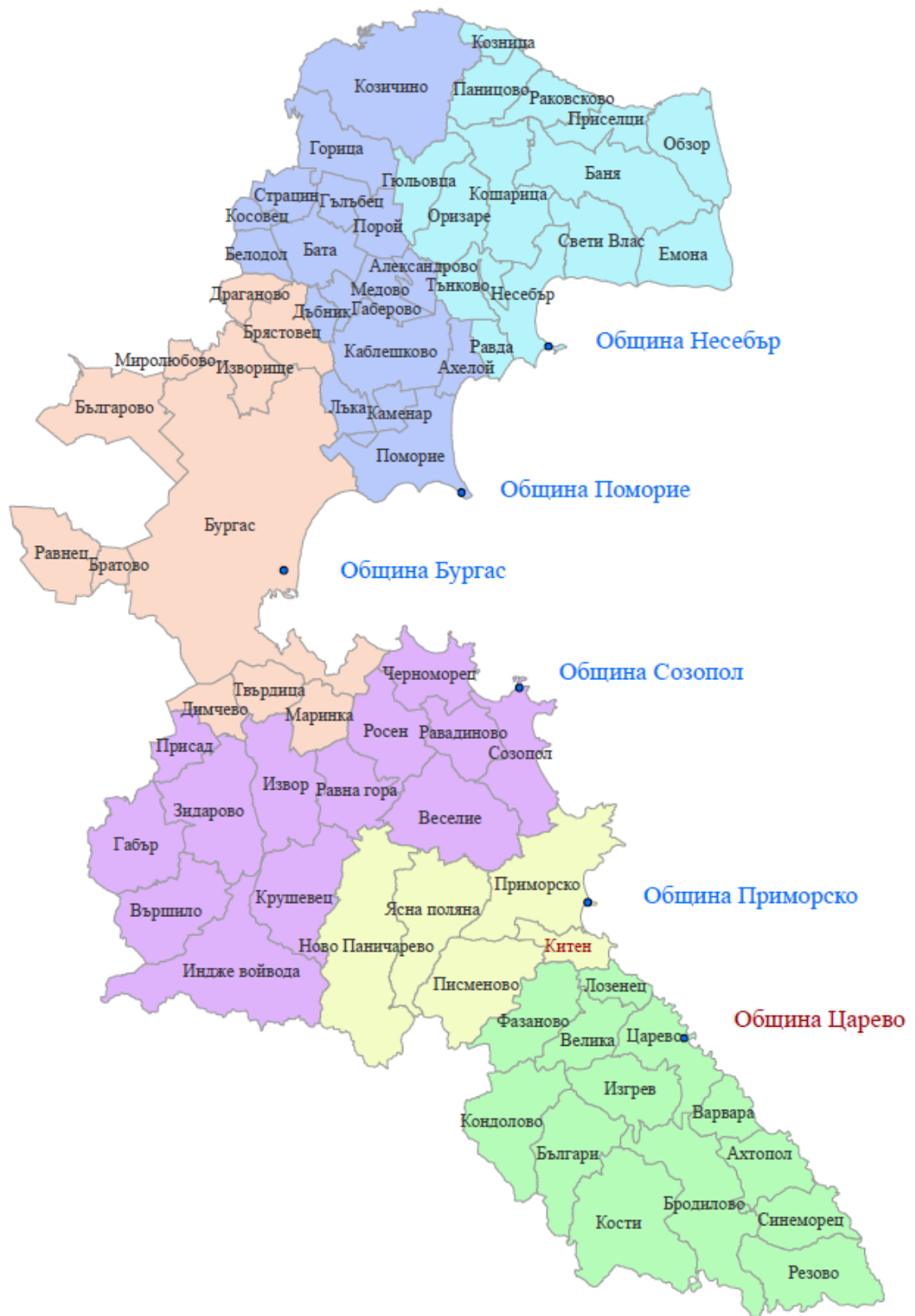
Източник: НСИ.

Фиг. 3. Населени места в състава на общините по Южното Черноморие на България към 12.07.1997 година



Източник: НСИ.

Фиг. 4. Населени места в състава на общините по Южното Черноморие на България към 18.04.2001 година



Източник: НСИ.

Фиг. 5. Населени места в състава на общините по Южното Черноморие на България към 25.04.2002 година



Източник: НСИ.

Фиг. 6. Населени места в състава на общините по Южното Черноморие на България към 22.05.2010 година



Източник: НСИ.

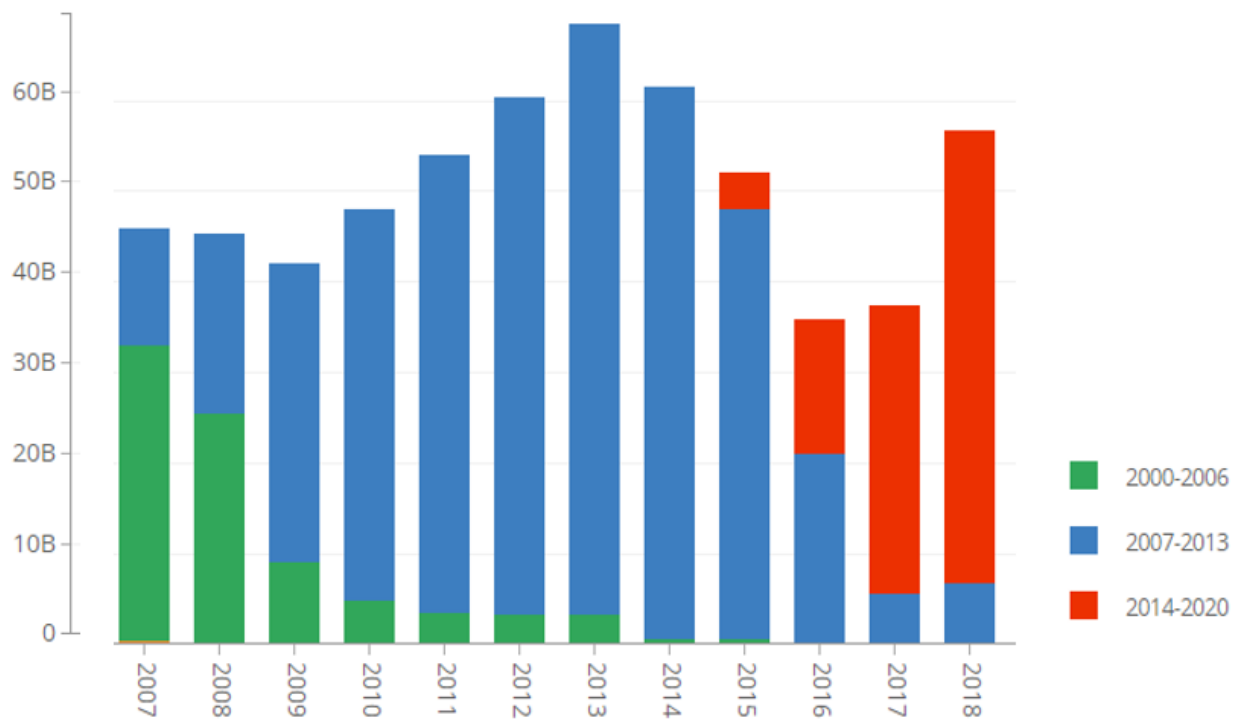
Част от поставените предизвикателства са адресирани в по-ранно изследване, в което предлагам в бъдеще административно-териториалното устройство на България и другите европейски държави да се основава на функционални райони от типа на районите на пазара на труда (РПТ), с цел ограничаване на ефекта, който исторически унаследените административни граници оказват върху качеството на данните на субнационално ниво⁸. Освен това едно такова функционално райониране, обективно отразяващо социално-икономическата реалност, би било много по-стабилно във времето в сравнение с честите и понякога субективни административно-териториални и териториални промени.

2. За необходимостта на Европейската комисия от по-дълги динамични редове от регионални данни и други субнационални данни

Програмите, свързани с кохезионната политика на ЕС, са с продължителност 7 години, но действителната им „продължителност на живота“ е по-дълга. Това може най-лесно да се обясни посредством следващите 2 фигури, представящи размера на годишните плащания от бюджета на ЕС, реализирани по програмни периоди общо за Европа и към най-бедния NUTS 2 регион на България, а именно Северозападния. Програмните периоди са избрани така, че да съвпадат с началото на членството на България в ЕС.

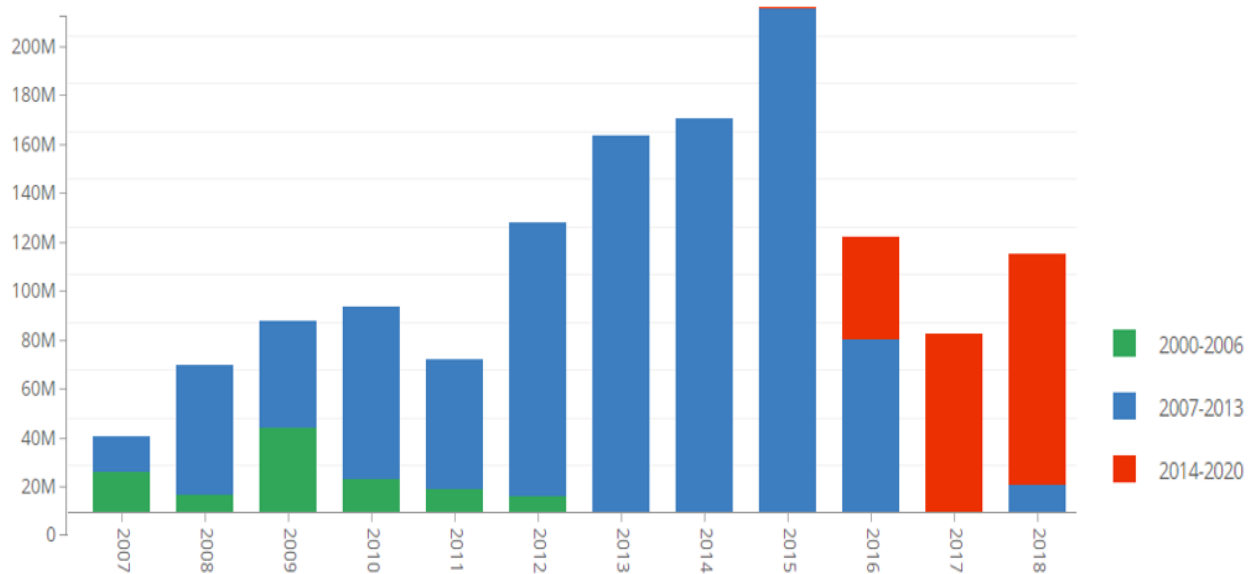
⁸ Ангелова, В. (2021). Методи и концепции за функционално географско райониране за статистически цели, Статистика, бр. 1/2021 https://www.nsi.bg/spisaniestatistika/page/bg/details.php?article_id=255

Фиг. 7. Годишни плащания от бюджета на ЕС, реализирани по програмни периоди общо за Европа (милиарди евро)



Източник: <https://cohesiondata.ec.europa.eu/EU-Level/Annual-EU-budget-payments-EUR-made-by-programme-pe/h85u-sjcv>. Селекцията е направена по димензиите „общо за всички фондове“ и „общо за ЕС“.

Фиг. 8. Годишни плащания от бюджета на ЕС към NUTS 2 - „Северозападен“, реализирани по програмни периоди (милиони евро)



Източник: <https://cohesiondata.ec.europa.eu/EU-Level/Annual-EU-budget-payments-EUR-made-by-programme-pe/h85u-sjcv>. Селекцията е направена по дименсиите „общо за всички фондове“ и „NUTS 2 – „Северозападен“.

Категориите региони (по-развити, региони в преход и по-слабо развити региони) са законово определени за целия 7-годишен период и се основават на фиксирана версия на NUTS. Това означава, че при евентуална значителна промяна в границите на NUTS регионите данните, на които се основават кохезионната политика към някакъв момент в рамките на програмния период, вече няма да са актуални. Мониторингът на програмите изисква адекватни социално-икономически данни с консистентни времеви редове. Оценката на дългосрочното въздействие на териториалните политики се нуждае от анализи, които надхвърлят времевата рамка на един програмен период. Създаването на доклад за социално-икономическото сближаване в Европа на всеки три години е законово задължение на ЕК и по-специално на DG REGIO (Directorate-General for Regional and Urban Policy). Това задължение е уредено в чл. 175 от Договора за функционирането на Европейския съюз, част трета „Политики и вътрешни действия на Съюза“, Дял XVIII - Икономическа, социална и териториална кохезия“⁹. Периодите, за които се отнасят тези доклади няма как да бъдат напълно съчетани с датите на влизане в сила на нова версия на NUTS или със сроковете на предвидените задължения на

⁹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12016E175>

държавите членки по предаване на преизчислението на данните. От друга страна, макар че преизчислението на динамичните редове спрямо предходни версии на NUTS е регламентирано, както вече беше обяснено, по различните статистически теми не се изисква предоставяне на еднаква дължина на ревизираните редове от данни. Как тогава да се направи оценка на въздействието на териториалните политики между два програмни периода за територии с промени в границите на NUTS и по показатели, за които предоставянето на преизчислени данни не е задължително? Този въпрос е подробно адресиран в следващата секция от анализа.

Нуждите на ЕК, свързани с по-дълги динамични редове и повече и по-детайлни статистически данни, съвсем не се изчерпват с гореизложеното. Фокусът на политиките на сближаване не се ограничава до териториалните измерения, въпреки че те са основополагащи, а се простира отвъд административните и дори държавните граници. Такива са политиките, съсредоточени върху околната среда, климатичните промени, пандемията и други. Това е неизменно свързано с фокус върху районите, обект на програми за трансгранично сътрудничество и действия, насочени към подпомагане на партньорството между градските и селските територии, както и съответните инструменти за тяхното подпомагане. Следователно докладите относно имплементирането на такива програми изисква разнообразни и своевременно актуализирани субнационални данни с достатъчно дълги времеви редове независимо от периодично възникващите промени в границите на териториалните и административно-териториалните единици.

Така нареченият Регламент TERCET, който допълва Регламента за NUTS, е публикуван през декември 2017 година¹⁰ посредством този законодателен акт различни териториални типологии като тези на градските и селските региони, на метрополните региони и на крайбрежните региони (базирани на NUTS 3), степента на урбанизация, градовете и техните функционални градски райони (ФГР) (от английски - Functional Urban Areas (FUAs), крайбрежните райони и степента на урбанизация (базирани на LAU) са включени в консолидираната версия на Регламента за NUTS. Регламентът за изпълнение на TERCET, публикуван през 2019 г., задава единните условия за хармонизирано прилагане на териториалните типологии¹¹. Идеята за наличието на хармонизирани европейски дефиниции по териториалните типологии е да се гарантира

¹⁰ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32017R2391>

¹¹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019R1130>

устойчивост на резултатите и международна сравнимост на статистическите данни. От своя страна обаче, всяка промяна в границите на NUTS 3 регионите и на местните административни единици автоматично налага своевременна актуализация на съответните класове на териториалните типологии. В този случай отново възниква предизвикателство, свързано с преизчислението на данни, агрегирани по тези типологии като преизчислението на данни по типологии, базирани на NUTS 3, което засяга в по-голяма степен Евростат отколкото националните статистически институти, тъй като данните биват автоматично агрегирани, респективно преагрегирани в производствената система на Евростат на базата на входни NUTS 3 данни, подадени от държавите членки¹². При сегашната практика за потребителите това също означава прекъсване на динамичния ред. В такъв случай Евростат обозначава данните за първата референтна година, в която настъпилата териториална промяна дава отражение върху динамичния ред с „b“ (break in time series), а клетките, за които има налични преизчислени данни - с „r“ (revised). Практиката в НСИ е да се поставя обяснителна бележка, съпровождаща таблиците, която би могла да остане незабелязана от потребителите на данни.

Най-търсените от потребителите данни по типологии, базирани на местни териториални единици, безспорно са степента на урбанизация (задължителна променлива във всички социални изследвания), градовете и техните функционални градски райони (ФГР). Ежегодната актуализация на степента на урбанизация се налага от често настъпващите промени в границите на LAU. Тъй като градовете, или т. нар. в Регламент TERSET „гъсто населени райони“, са една от категориите, включени в степента на урбанизация, като всяка настъпила промяна в класовете по степен на урбанизация налага корекции в списъка с градовете. От 2017 г. насам държавите, които имат задължение да изпращат на Евростат ежегодните списъци с LAU, доброволно предоставят допълнителна информация относно текущата типологизация на местните административни единици. Тази информация става официална едва след валидиране от страна на Евростат, като по технически причини комуникацията понякога отнема месеци. Често срещана в практиката ситуация е, че Евростат работи с последния официален грид на населението от 2011 г. и с границите на LAU, които организацията EuroGeographics ежегодно произвеждат за ЕК¹³, докато някои държави членки

¹² НСИ понастоящем не публикува на интернет страницата си данни, разпределени по териториалните типологии, включени в TERSET. Наличните данни по град - село се основават на дефинициите според националното законодателство.

¹³ <https://eurogeographics.org/maps-for-europe/ebm/>

разполагат с по-детайлен слой на границите и с по-актуален грид на населението понякога дори с различна резолюция. Това води до разминаване в получените резултати по класове типологии и в тези случаи Евростат възприема това, което страните са предоставили. След постигане на съгласие относно класификациите по степен на урбанизация и останалите типологии, основани на LAU, Евростат и националните статистически институти предоставят информацията на тематичните отдели в съответните структури, които използват териториални типологии за производство на данни за градовете, туризма и данни от социални изследвания като Наблюдението на работната сила, Изследването за доходите и условията на живот, Наблюдението на домакинските бюджети, Изследването на образованието и обучението на възрастни и редица други. Една от най-комплицираните методологии за ежегодна актуализация на LAU-базираните териториални типологии е степента на урбанизация. Тя обаче за пореден път демонстрира мощта на грида като помощен слой за функционално географско райониране, респективно за производство на данни¹⁴.

На семинар под наслов “Териториалната кохезия след 2020: Интегрирано териториално развитие за по-добри политики”, организиран от ESPON (European Spatial Planning Observation Network) по случай Българското председателство на Съвета на ЕС, проведен в София през 2018 г., Барка¹⁵ споделя с аудиторията, виждането си че: „ЕС спешно се нуждае от масово и убедително прилагане на подход, базиран на локализацията... За тази цел, последното нещо, което трябва да се направи, е да започнат преговори за периода след Бюджет 2020, в който ... интегрираният подход за развитие на градските и селските територии се третира просто като поредния сектор или ниша, вместо като хоризонтална характеристика на цялата стратегия за териториално сближаване." Можем да твърдим, че териториалните типологии, създадени за статистически и аналитични цели, отговарят на философския смисъл, вложен в изказването на проф. Барка. Между различните типологии съществуват тесни връзки, основаващи се на местоположението, което дава възможност за интегрирано ползване и интерпретация на съответните статистически данни. Например крайбрежен район може

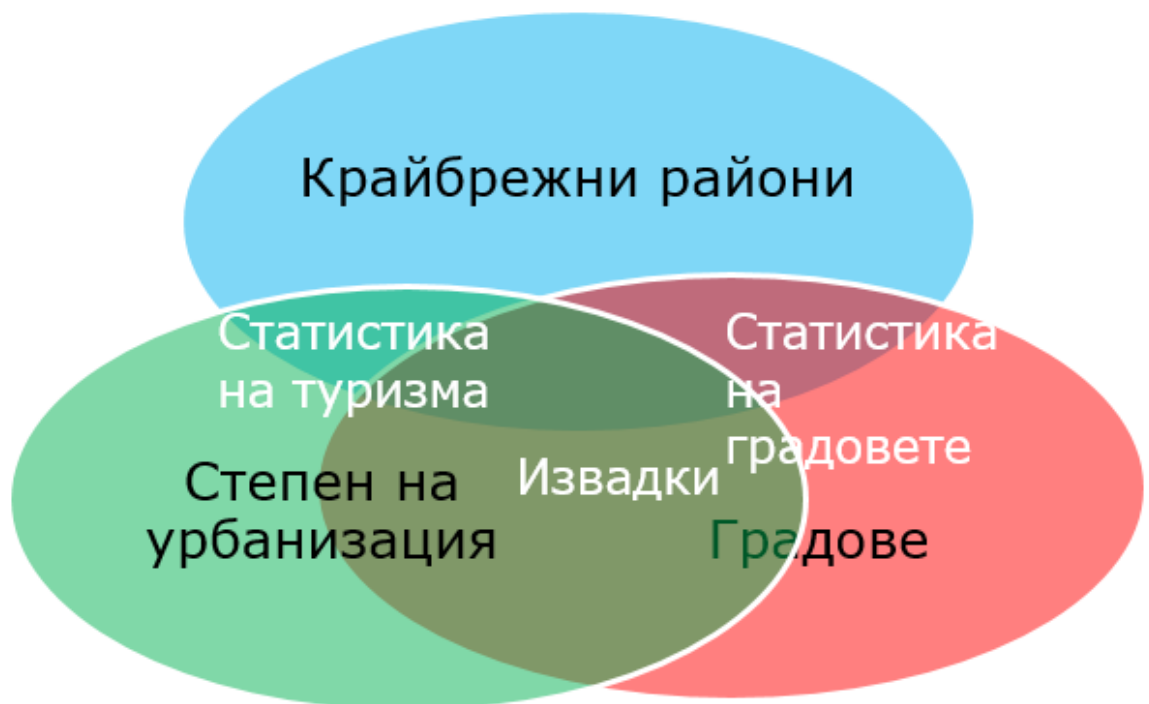
¹⁴ [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Territorial typologies manual - degree of urbanisation#Changes to the typology over time](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Territorial_typologies_manual_-_degree_of_urbanisation#Changes_to_the_typology_over_time) Информацията може да се намери в секция „Актуализация на степента на урбанизация за отразяване на промените в границите на LAU“ (Updating the degree of urbanisation to reflect changes in LAU boundaries).

¹⁵ Проф. Фабрицио Барка е италиански икономист и бивш министър без портфейл, отговарящ за териториалното развитие, в експертния кабинет на Марио Монти (2011 - 2013 г.). Към август 2021 г. презентацията е достъпна на адрес <https://www.espon.eu/sofia>

същевременно да е селски (в България такива са общините Аврен, Балчик, Бяла, Долни чифлик, Несебър, Поморие, Приморско, Созопол и Шабла) или да спада към категорията на по-малките градове и предградия (Аксаково, Каварна и Царево) или пък да е градски център в границите на метрополен регион (Бургас и Варна), град София е разположен в планински регион, а в състава на неговия ФГР има и селски райони, тъй като от тях произлизат значителни потоци от ежедневни трудови мигранти към столицата и т.н.

На фиг. 9 са представени схематично авторски разсъждения относно имплементирането на LAU списъците в официалната статистика и връзката между различни типологии и статистически изследвания.

Фиг. 9. Схема на имплементирането на LAU списъците в официалната статистика



Единственото готово към момента практическо решение на проблемите, описани дотук, е т. нар. NUTS конвертор, разработен от Joint Research Centre (JRC) към Европейската комисия (ЕК)¹⁶. Този инструмент бива създаден и се поддържа с цел

¹⁶ <https://urban.jrc.ec.europa.eu/nutsconverter/#/>

подпомагане на потребителите при работата им със статистическите данни, докладвани по различни версии на NUTS. NUTS конверторът е свободно достъпен уеббазиран инструмент, позволяващ преобразуването на регионални данни между различни версии на NUTS класификацията, докладвани на нива по NUTS 1, 2 и 3, преизчислява данни между версиите на NUTS от 2006, 2010, 2013 и 2016 г. във всяка посока, като автоматично разпознава **конкретната** версия, на която се основават входните данни. Конверторът се основава на набор от изчислени матрици за търсене, установяващи пространствената връзка между региони от различни версии на NUTS, и трансформациите, които те са претърпели в резултат от прерайонирането. Преизчисленията се извършват на базата на пространствена информация от пет помощни географски слоя с висока разделителна способност (площ на регионите, застроени площи, брой жители според грида на населението, дневно и нощно население). Потребителят има възможност сам да избере помощната променлива, по която да се извършат изчисленията. Инструментът борави с всички региони на NUTS в рамките на 27-те държави членки на Европейския съюз (ЕС-27), държавите - членки на Европейската асоциация за свободна търговия (ЕАСТ) и Обединеното кралство. Въпреки че инструментът е базиран на строги научни методи за пространствена корелация, той има някои технически недостатъци (повечето от данните по помощните променливи все още са базирани на Преброяване 2011). Инструментът не е проектиран да замени официалната регионална статистика и ЕК не поема отговорност за резултатите от използването на инструмента. Следва да направим и едно уточнение относно възможностите на ЕК в лицето на Евростат да преизчислява данни и то по-специално от извадкови изследвания. Националните статистически институти включват кодовете на NUTS 2 и 3 регионите по местоживееене на домакинствата (в Наблюдението на работната сила разполагаме дори с променлива „регион по месторабота“) в микро-данните, които предоставят на Евростат, но не са задължени да дават индикация за кодовете на местните административни единици. Оттук следва, че за Евростат е практически невъзможно да произвежда оценки на данни за малки териториални единици, каквито са градовете и техните ФГР, и тази задача остава изцяло в ръцете на националните производители на данни.

3. Гео реферирането като решение на широк набор от проблеми, стоящи пред националната и европейската статистическа теория и практика

Авторът вижда гео кодирането като оптимално решение на многобройните предизвикателства, изложени вече, като дефинира и някои допълнителни приложения на това сравнително ново направление в традиционната статистическа практика на нашата страна.

На европейско равнище необходимостта от гео статистиката за целите на преброяването на населението и жилищния фонд е всепризната. Регламентът за изпълнение (ЕС) 2018/1799 на Комисията от 21 ноември 2018 г. за започване на временно пряко статистическо действие за разпространение на избрани теми от Преброяването на населението и жилищния фонд през 2021 г., гео кодирани към координатна мрежа от 1 кв. км, е първият по рода си в Европейската статистическа система¹⁷. В Регламента е записано, че „това действие е обосновано от общата нужда в целия Съюз от надеждна, точна и съпоставима информация за разпределението на населението с достатъчна пространствена разделителна способност въз основа на хармонизирани изисквания към резултатите и предназначена по-конкретно за общоевропейската регионална политика“. Избраните теми, които следва да бъдат гео реферирани, са: обичайно местопребиваване, пол, възраст, статут на настояща дейност (брой наети лица), държава/място на раждане и обичайно местопребиваване една година преди преброяването. Никак не е изненадващ фактът, че при гео реферирането ще се сблъскаме с клетки с нулев или близък до нулата брой население по различните разбивки. Редица клетки са покрити от природни обекти като езера, големи реки, обработваеми земи, високопланински територии и други. Въпреки съществуващите ограничения считам грида за универсален географски компонент за производство на данни, за гео пространствени анализи и дефиниране на функционални географски райони. Освен това Регламентът урежда един съществен въпрос, свързан с конфиденциалността на данните, осигурявайки възможност на държавите членки да спазят и националното си законодателство¹⁸. Стратегията на Евростат за развитието на демографската статистика след Преброяване 2021 предвижда

¹⁷ https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=uriserv%3A0J.L_.2018.296.01.0019.01.BUL&toc=OJ%3AL%3A2018%3A296%3ATOC

¹⁸ Член 2, ал. 8 от Регламента дава следното определение: „поверителна стойност“ означава цифрова стойност, която не трябва да бъде разкривана с цел запазване на статистическата поверителност на данните и в съответствие със защитните мерки на държавите членки срещу оповестяването на статистически данни. Член 5, ал. 2 определя, че държавите членки заменят всяка поверителна стойност със специалната стойност „Не е известно“.

държавите членки ежегодно да докладват гео кодирани данни за броя на населението според обичайното местожителство¹⁹. Очакванията са ежегоден грид на населението да започне да се поддържа и разпространява през 2024 година. Универсалният характер на грида може да бъде илюстриран и чрез следния пример за възможност за бъдеща изследователска работа с данните от предстоящото Преброяване и данни от последващи актуализации на грида на населението, а именно анализ на положителния/отрицателния прираст на населението по грид клетки, респективно по клъстери от грид клетки.

Необходимостта от грида като детайлно географско ниво за събиране и анализ на статистическа информация се потвърждава и от Регламент (ЕС) 2018/1091 на Европейския парламент и на Съвета от 18 юли 2018 година относно интегрирана статистика за земеделските стопанства и за отмяна на регламенти (ЕО) № 1166/2008 и (ЕС) № 1337/2011²⁰. Този нормативен акт урежда събирането на нова променлива - код на клетката от координатната мрежа, в която се намира земеделското стопанство.

Концепцията за грида, разбира се, е включена и в Регламент TERSET, тъй като тя е основата за производство на териториалните типологии. Единствено типологията на крайбрежните райони не се базира на грида на населението, а на наличието на морска граница или изпълнено условие за поне 50% от площта на местната административна единица, намиращи се на разстояние в рамките на 10 км от крайбрежието. Споменатите законодателни актове и стратегията дават ясна индикация, че гео кодирането на избрани теми от демографската и земеделската статистика е само началото на пътя, който следва да извървим към постепенното гео рефериране на всички официални статистически изследвания - изчерпателни, извадкови и такива, базирани на регистри.

Основният въпрос, на който настоящото изследване се стреми да отговори, е как посредством гео кодирането ще осигурим достатъчно гъвкав процес на събиране на данни, при който да може адекватно да се реагира на сравнително честите промени в границите на териториалните и административно-териториалните единици, осигурявайки в същото време достатъчно дълги динамични редове и данни на различни териториални нива, включително такива по териториални типологии, както и данни,

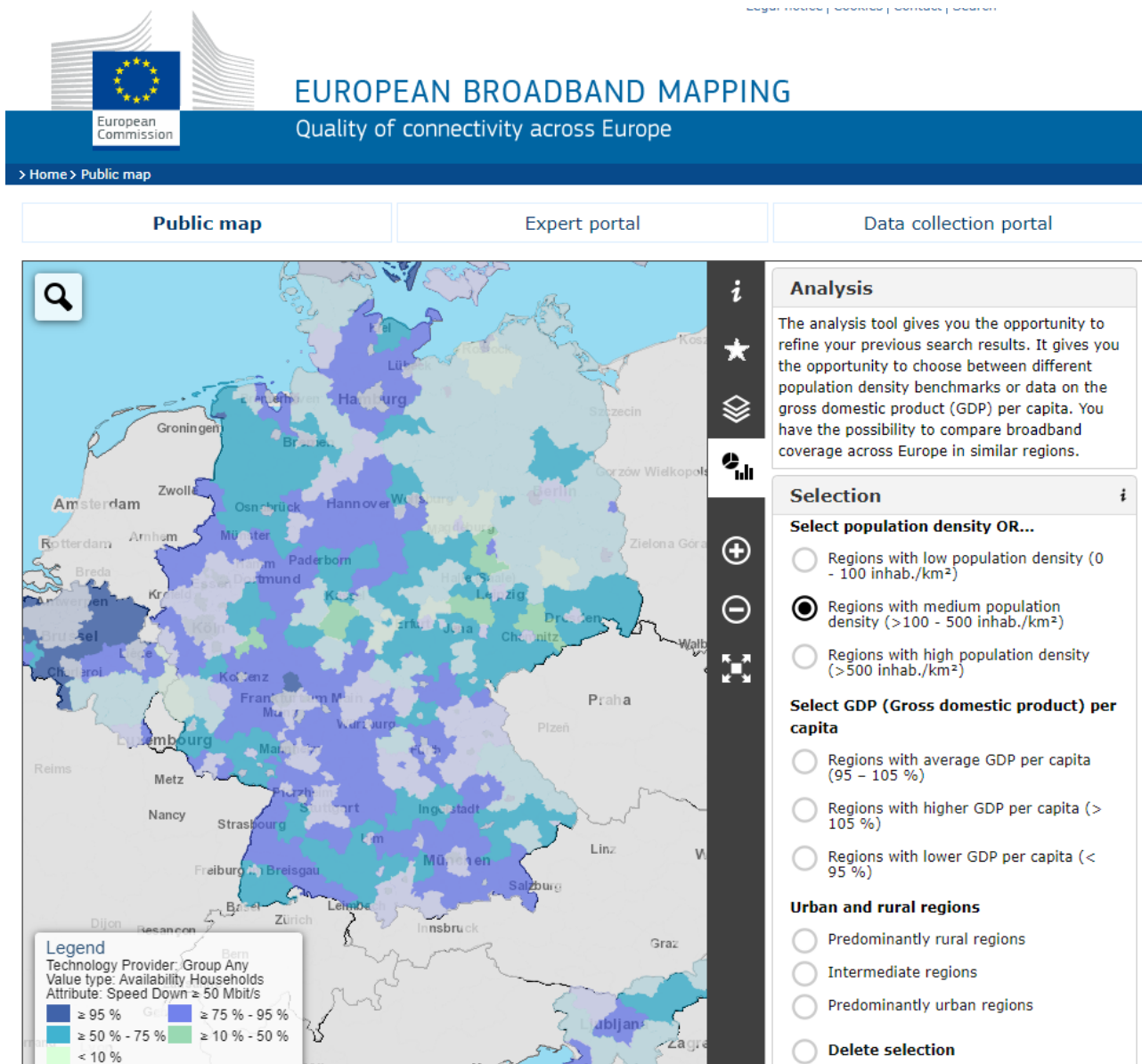
¹⁹ <https://circabc.europa.eu/sd/a/e9072c55-f448-47f6-b4e2-cb3697ae7dc1/DSS-2016-Oct-3.2%20Census%20post%202021%20vision.pdf> (за достъп до съдържанието се изисква регистрация на страницата)

²⁰ https://eur-lex.europa.eu/legal-content/BG/TXT/?uri=uriserv%3AOJ.L_.2018.200.01.0001.01.BUL&toc=OJ%3AL%3A2018%3A200%3ATOC

произхождащи от извадкови изследвания? Този механизъм следва да осигурява навременния мониторинг на провежданите в България и Европа териториални политики.

Нека първо допълнително да илюстрирам аналитичната мощ на грида в подкрепа на вземането на мъдри решения по въпросите на секторните и териториалните политики на ЕС и политиката на сближаване. През 2017 г. генералната дирекция „Комуникационни мрежи, съдържание и технологии“ (DG CONNECT) осъществява проект за създаване на портал за интерактивно картографиране на качеството на интернет мрежата в Европа по отношение на широколентовия достъп. В процеса на подготовка на портала заинтересованите страни в лицето на публични и частни доставчици разработват стандарти за обобщаване на данните, като над 30% от тях са предоставили исканите данни за широколентов достъп до интернет. Впоследствие DG CONNECT съвместно с Евростат доизграждат прилежащата географска информация, необходима за визията и съдържанието на портала. Интерактивната картографска платформа позволява на потребителя да направи селекция по различни характеристики на регионите и скоростта на интернет връзката в Европа. На фиг. 10 е представена стандартна регионална карта, позиционирана според наличността на данните на европейско равнище, като показаната селекция съчетава гъстота на населението и наличието на широколентова интернет връзка със скорост по-голяма или равна на 50 мегабита в секунда (Mbit/s). В случая корелацията между социално-икономическото развитие на показаните региони и качеството на широколентовите услуги е чисто интуитивна поради липса на входните данни за по-задълбочени изчисления.

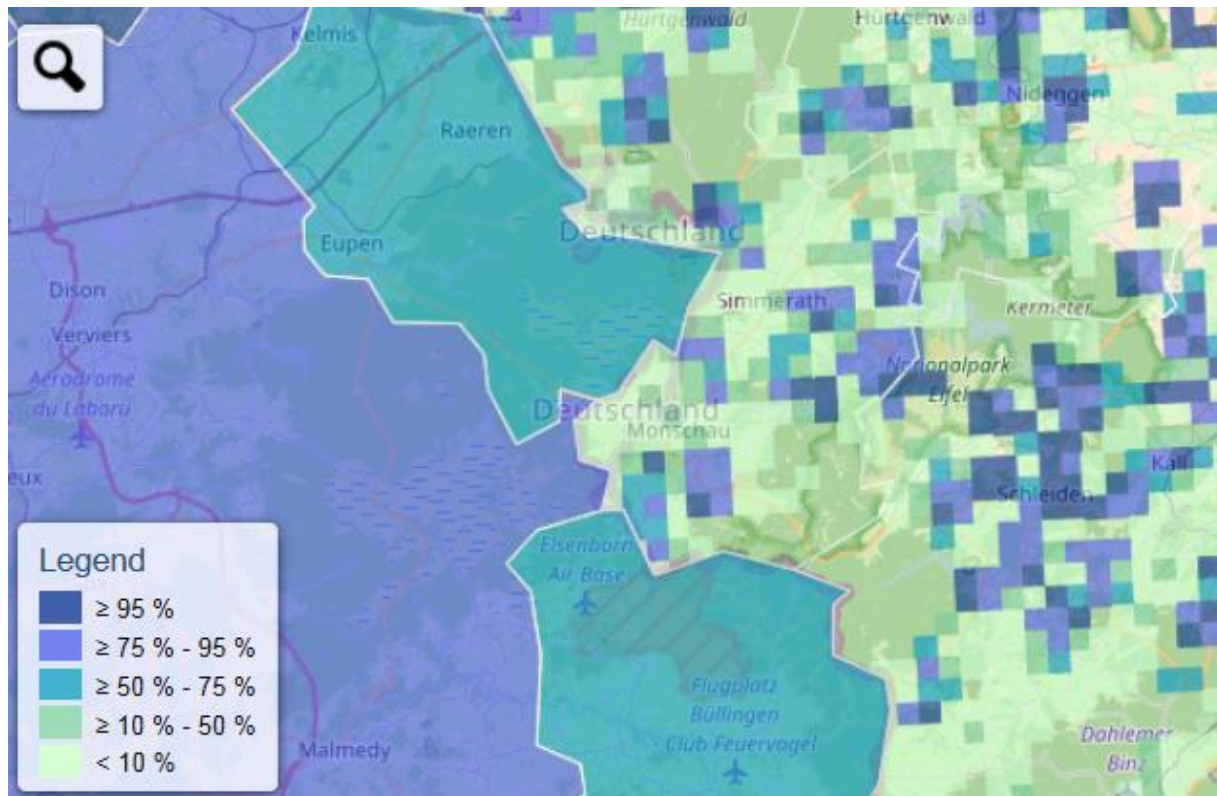
Фиг. 10. Относителен дял на домакинствата в избрани региони с гъстота на населението между 100 и 500 души на 1 кв. км, разполагащи с широколентова интернет връзка със скорост по-голяма или равна на 50 мегабита в секунда



Фигура 11 представя относителния дял на домакинствата, които имат достъп до широколентова интернет връзка в мрежа със скорост от 50 мегабита в секунда. Данните за Белгия са налични на ниво NUTS, докато тези за Германия - на ниво грид клетки със страна от 1 квадратен километър. Изображението е създадено експериментално в рамките на проекта, за да покаже както на вземащите политически решения, така и на частните инвеститори в интернет мрежи как данните на ниво грид дават далеч по-добра картина на достъпа на населението до тази услуга в сравнение с традиционните регионални данни на ниво NUTS. Тъй като платформата не е еднократна инициатива на

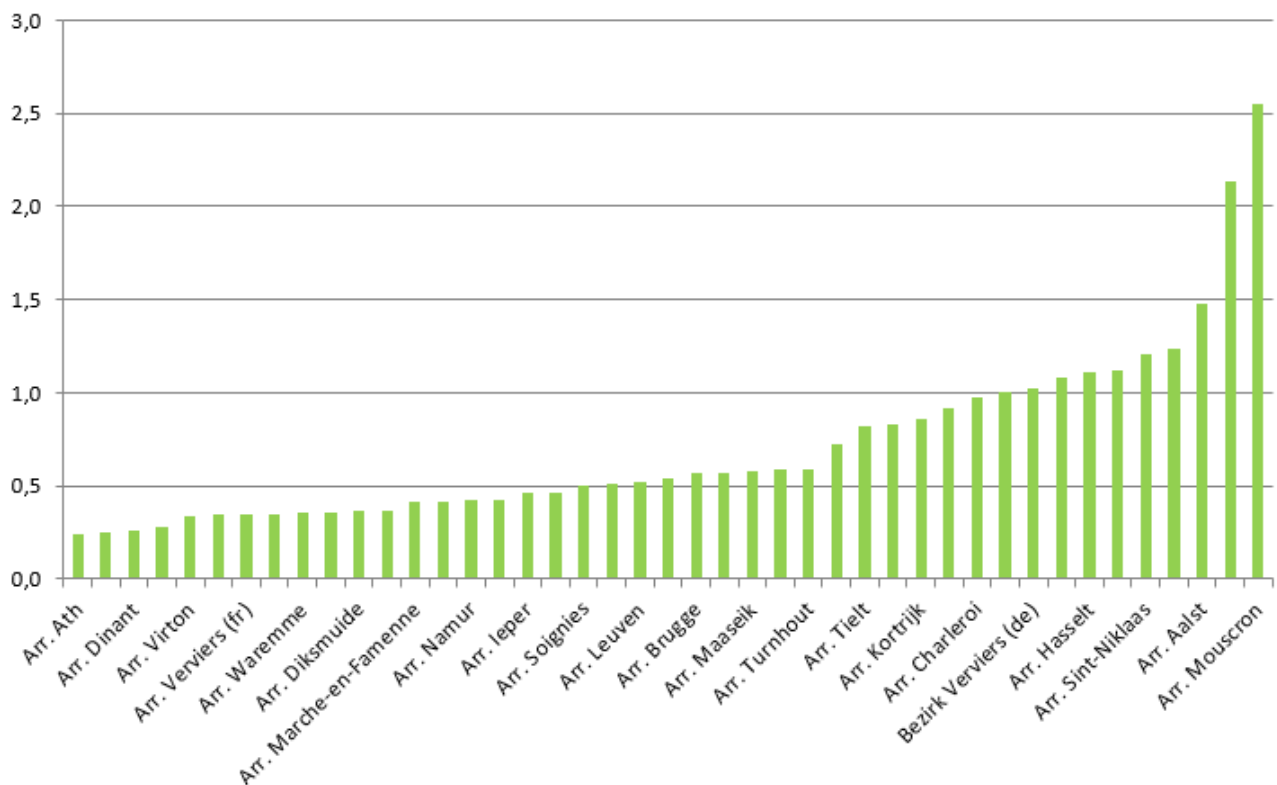
ЕК, за в бъдеще тя ще предоставя възможност на заинтересованите страни да наблюдават напредъка, постигнат в териториален аспект по отношение на разгръщането на интернет мрежи с голям капацитет и качеството на широколентовите услуги в Европа.

Фиг. 11. Относителен дял на домакинствата, които имат достъп до широколентова интернет връзка със скорост от 50 мегабита в секунда



Важно е да се отбележи, че за разлика от планираното и вече постигнатото в областта на демографската статистика и проекта на DG CONNECT гео кодирането на микроданните от извадките изследвания далеч не цели разпространението на данни на ниво грид клетка поради обективно съществуващите ограничения, произлизащи от поверителността на данните. Невъзможността за представяне на данни от извадкови изследвания на ниво грид клетка са илюстрирани на фиг. 10, която представя средногодишния брой на респондентите, включени в Наблюдението на работната сила, които се падат на обитавана грид клетка с площ 1 кв. км в NUTS 3 регионите на Белгия. Регионът на Брюксел със средно цели 34 460 респонденти на обитавана грид клетка не е включен в графиката, за да не се наруши нейната читаемост.

Фиг. 12. Гъстота на респондентите в Наблюдението на работната сила в NUTS 3 регионите на Белгия, 2017 година



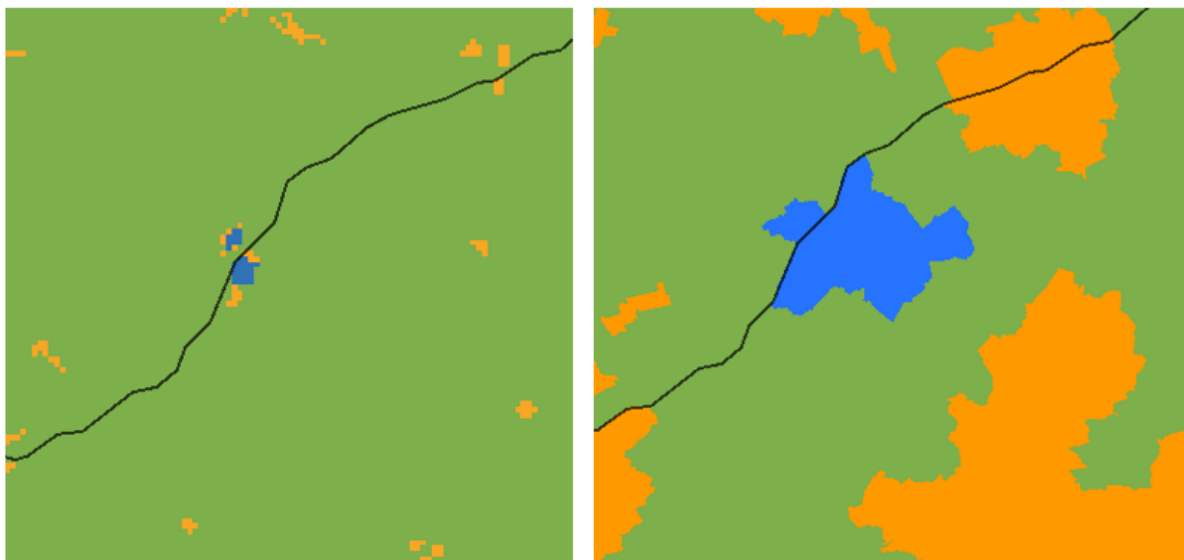
През 2022 г. се очаква да има обработени и валидирани резултати от предстоящото към момента на писане на статията преброяване, което означава, че НСИ скоро ще разполага не само с координатите на сградите в България, но и с координатите на всички потенциални респонденти, които биха могли да попаднат в извадките на социалните изследвания. Това ще отвори неограничени възможности за гъвкаво агрегиране на данни на ниво гريد клетка от социални изследвания, до която и да е версия на границите на LAU и NUTS. Добавянето на нова променлива „географски координати“ на сградата, в която респондентът живее, или „идентификационен номер на гريد клетката“, или „координати на центроида на гريد клетката“ в дизайна на извадките ще създаде възможност за преизчисление на данни до всевъзможни територии и функционални географски райони от потребителски/изследователски интерес (като например РПТ, споменати в точка 1, чиито конфигурации не са непременно имплементирани в дизайна на съответната извадка, както и за интегрирането на данните от изчерпателни, извадкови изследвания и данни от регистри въз основа на местоположението като ключ. Изборът относно това коя от обсъжданите по-горе променливи да бъде включена в дизайна на извадките следва да бъде обект на експертно

методологично решение. Във всеки случай този GRID базиран подход ще позволи значително съкращаване на времето, нужно за преизчисление на динамичните редове и за тяхното агрегиране по различни територии от интерес, като по този начин ще намали натовареността на производителите на данни. В случай на бъдеща необходимост гео кодирането на социалните изследвания ще позволи и преизчисление на самите GRID базирани данни, например посредством симулации.

Съществува практически проблем, който се нуждае от изграждане на персонализиран методологичен подход в случаите, когато задачата е да се произведат данни от извадкови изследвания на ниво функционални райони каквито са ФГР и РТП. Предварителните очаквания на автора са, че ще се налага допълнително претегляне на извадката.

Посредством гео кодирането може да се отговори не само на конкретни възникнали нужди и персонализирани потребителски заявки. В дългосрочен план качеството на данните може да бъде подобро в няколко направления. На първо място евентуално кодиране на респондентите по тип на клъстера според дефинициите и терминологията, заложи в Регламент TERSET („клетки от селски тип“, или „слабо населени клетки“, „клъстери от градски тип“, или „средно населени клъстери“ и „градски центрове“, или „много населени клъстери“), вместо кодирането по степен на урбанизация би минимизирало затрудненията, произтичащи от ежегодната актуализация на класовете по степен на урбанизация вследствие на промените в границите на местните административни единици. Последващото агрегиране на данните по тип на клъстера няма да представлява проблем, а прецизността на получените резултати ще е по-висока и прекъсването в динамичните редове ще е с по-малка честота. Степента на урбанизация, разбира се, може да се запази като алтернативна променлива във въпросниците на социалните изследвания, което ще позволи сравнение на резултатите и ще осигури различен мащаб за решаване на различни познавателни задачи. Различията между обхвата на клъстерите и класовете по степен на урбанизация са показани на фиг. 13 по примера на трансграничния район на Русе - Гюргево.

Фиг. 13. Клъстери от гريد клетки според гъстотата на населението (отляво) и степен на урбанизация (отдясно) в трансграничния район на Русе - Гюргево²¹



Източник: Статистически атлас на Евростат с данни за грида на населението от Преброяване 2011, предоставени от държавите членки.

Още една дългосрочна полза, която статистическата теория и практика могат да извлекат от наличието на гريد базирани данни, е възможността за оптимизиране на самите извадки на социалните изследвания. Експерти от националния статистически институт на Португалия вече са тествали алгоритъм, използващ крива за избор на обитавани гريد клетки във всяка страта на извадката с цел оптимизиране на извадката и по-добро покритие на територията. Сантос представя резултатите от това изследване на конференция на Европейския форум по гео статистика, проведена в София през 2013 година²². Допълнителна възможност, която би могла да се тества, е стратификация на извадките по тип на клъстерите според гъстотата на населението вместо по националната дефиниция за град и село, като стремежът е това да не окаже влияние върху обема на извадката.

²¹ <https://ec.europa.eu/eurostat/statistical-atlas/gis/viewer/?mids=BKGCNT,TYPC11,CNTOVL&o=1,1,0.7&ch=C01,TRC,TYP¢er=43.74753,26.45999,7&lcis=TYPC11&> <https://ec.europa.eu/eurostat/statistical-atlas/gis/viewer/?mids=BKGCNT,TYPU11,CNTOVL&o=1,1,0.7&ch=C01,TRC,TYP¢er=43.74753,26.45999,7&lcis=TYPU11&> В синьо са показани най-гъсто населените територии според всяка от класификациите, в оранжево - междинните територии, а в зелено - най-слабо населените територии.

²² Santos, A., B. J. Schoenmakers (2013). Using the European Grid 'ETRS89/LAEA_PT_1K' as the foundation for the new Portuguese Sampling Infrastructure, достъпен към август 2021 година, http://www.nsi.bg/efgs2013/data/uploads/presentations/DAY2_WS2_6_Presentation_SANTOS_ok.pdf

Друг аспект на разискваната проблематика е, че през последните няколко години глобалните климатични промени стават все по-осезаеми. Ако разполагаме с грид данни за броя на възрастното население, непълнолетните, лицата, лишени от трудоспособност, броя на лицата в риск от бедност и т.н., в случай на бедствия или аварии ГИС експертите от националните статистически институти своевременно ще могат да локализируют най-рисковите групи от населението, което ще подпомогне съответните служби за вземането на бързи и обосновани решения относно организацията на спасителните акции и дейностите по евакуация на населението.

За да стане картината още по-пъстра, ще добавя един последен (но не по важност) щрих. През 2013 г. НСИ на България разработи универсална методология за приложение на SAE, която съчетава принципи от клъстерния и структурния анализ с елементи от теорията на вероятностите. Така конструираната от Ангелова и Гавазки методология²³ позволява производство на данни от извадкови изследвания на ниво град, ФГР и стига дори до ниво населено място. Най-общо казано, идеята се базира на създаването на клъстери от хомогенни в демографското и социалното си развитие клъстери от населени места и на хипотезата, че в населени места от един и същ клъстер вероятността за проявление на даден феномен е еднаква. Счита се, че във всички населени места от един клъстер заетостта, броят на рано напусналите образователната система, броят на студентите, хората, живеещи под линията на бедност, и т.н. биха били еднакви. Актуализацията на основата на тази методология предстои като входни данни на ниво грид вместо по населени места и би могла да се окаже възможност за още една иновация в официалната статистика.

Заклучение

В заключение си позволявам да отправя следните предложения и да направя някои констатации, които имат за цел една още по-гъвкава национална, а защо не и европейска статистика в услуга на потребителите.

Предлагам да се тестват възможностите на грида за оптимизиране на извадките на социалните изследвания на базата на данните от предстоящото преброяване на населението.

²³ Гавазки, И., В. Ангелова (2013). Методология за извършване на оценки от извадкови изследвания на ниво общини и населени места посредством клъстерен и структурен анализ, Статистика, бр. 1 - 2/2013 https://www.nsi.bg/spisaniestatistika/page/bg/details.php?article_id=33

Считам, че в бъдеще националните статистически институти ще могат да предложат нови услуги на потребителите - преизчисление на динамични редове за територии с променени граници, производство на данни по различни функционални географски райони, пространствени анализи, основани на GRID като например анализи за интелигентно локализиране на услуги и производства в подкрепа на бизнеса и редица други.

Препоръчвам въвеждане на единна система за обозначаване на данните, особено в случаите на ревизирани данни и прекъсване на динамичните редове по подобие на практиката на Евростат (flag „b“; „r“).

Намирам, че потенциалът на техниките за приложение на SAE не е използван в достатъчна степен. SAE в съчетание с GRID като универсално решение за интегриране на различни изчерпателни и извадкови изследвания и данни от регистри биха открили нови хоризонти пред официалната статистика.

Използвани съкращения

БВП	- брутен вътрешен продукт
ЕАСТ	- Европейска асоциация за свободна търговия
ЕК	- Европейска комисия
ЕКАТТЕ	- Единен класификатор на административно-териториалните и териториалните единици
ЕС	- Европейски съюз
НСИ	- Национален статистически институт
РПТ	- Райони на пазара на труда
ФГР	- Функционални градски райони
„b“	- Break in time series
DG REGIO	- Directorate-General for Regional and Urban Policy
ESPON	- European Spatial Planning Observation Network
Flag	- флагуване/обозначаване на данните с допълнителна информация
FUA	- Functional Urban Areas
JRC	- Joint Research Centre (съвместен Изследователски център към Европейската комисия)
LAU	- Local Administrative Unit (местна административна единица)
NUTS	- La nomenclature des unités territoriales statistiques
„r“	- Revised
SAE	- Small Area Estimation

ЦИТИРАНА ЛИТЕРАТУРА:

Ангелова, В. (2021). Методи и концепции за функционално географско райониране за статистически цели, Статистика, бр. 1/2021 https://www.nsi.bg/spisaniestatistika/page/bg/details.php?article_id=255

Гавазки, И., В. Ангелова (2013). Методология за извършване на оценки от извадкови изследвания на ниво общини и населени места посредством клъстерен и структурен анализ, Статистика“, бр. 1 - 2/2013 https://www.nsi.bg/spisaniestatistika/page/bg/details.php?article_id=33

Бояджиев, В., Н. Веселинова (2017) Границите като проблем за регионалното развитие на България. В: Сб. Юбилейна международна научна конференция БЪЛГАРИЯ НА РЕГИОНИТЕ, Перспективи за устойчиво регионално развитие, Пл.; <http://regions.uard.bg> (31.08.2021 г.; 15:00)

Eurostat (2019). Methodological manual on territorial typologies, Luxembourg <https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-18-008>

Eurostat, Collaboration in Research and Methodology for Official Statistics, Groups, General interest groups, Labour Market Areas, достъпен към август 2021 г., https://ec.europa.eu/eurostat/cros/content/geocoding-more-data-functional-areas-and-longer-time-series_en

Santos, A., B. J. Schoenmakers (2013). Using the European Grid ‘ETRS89/LAEA_PT_1K’ as the foundation for the new Portuguese Sampling Infrastructure, достъпен към август 2021 г., http://www.nsi.bg/efgs2013/data/uploads/presentations/DAY2_WS2_6_Presentation_SANTO_S_ok.pdf