From place of residence to place of activity: towards spatiotemporal mapping of population density in Europe

Filipe Batista e Silva*, Konštantín Rosina, Marcello Schiavina, Mario Alberto Marín Herrera, Sérgio Freire, Lukasz Ziemba, Massimo Craglia and Carlo Lavalle

European Commission Joint Research Centre, Ispra, Italy

*Corresponding author email address: filipe.batista@ec.europa.eu

Abstract

Current knowledge of the spatial distribution of population is still very incomplete. It is based upon place of residence statistics and does not account for the fact that people shift between various locations in daily, weekly and annual cycles for reasons of shelter, work, leisure or fulfilling various necessities. Spatial mobility of people results in significant variation of local population densities, which is extremely relevant for a range of applications, from risk assessment to urban and regional planning. Despite the advances in computational capacity and data availability, spatiotemporal mapping of population remains challenging and the state-of-the-art is considerably thin. ENACT ("ENhancing ACTivity and population mapping") is an ongoing applied research project aiming at producing the first EU-wide, consistent, seamless, high-resolution and validated population density grids that take into account major daily and seasonal population variations. This short paper provides an overview of the main tasks of the project, including: 1) estimation of regional and monthly stocks of population sub-groups; 2) detailed mapping of socioeconomic activities; 3) disaggregation of population stocks to grid level; 4) cross-comparison of the project's outputs with independent sources.
Keywords: Population density; Spatiotemporal mapping; Big Data; Data fusion; High-resolution; Europe.

1 Introduction

1.1 Background

Population is a crucial variable for the social sciences, the geosciences, and for policy support in many domains. Yet, our knowledge of its spatial distribution is still nowadays very incomplete. Population is a temporally dynamic variable, with major shifts in its distribution occurring in daily and seasonal cycles, resulting in rapidly changing densities. Spatially detailed representations of residential population exist for the European Union (EU) level since several years (Gallego *et al.*, 2011; Batista e Silva, Gallego & Lavalle, 2013). While these maps can be used as proxies for night-time population distribution, the distribution of population for other time frames is practically unknown at almost every spatial scale. Consequently, all applied sciences and policy support that require spatially detailed information on population distribution must rely on a fractional and static representation of reality. Overcoming this large knowledge gap is the main goal of the ENACT ("ENhancing ACTivity and population mapping") research project.

The location of population during the day is determined by the location of economic, social and leisure facilities which pull population off their residences, driving commuting flows and other forms of daily trips. Daytime population distribution thus varies greatly from night-time distribution, and it is much more challenging to infer.

Addressing the needs of emergency response, compatible day- and night-time population grids have been produced in the mid-2000s for the USA (McPherson & Brown, 2004; Bhaduri *et al.*, 2007). In Europe, such datasets have been mostly lacking, with only a few countries systematically collecting base data and modelling population distribution on the daily cycle (e.g. Ahola *et al.*, 2007). More recent research (Martin, Cockings & Leung, 2010, 2015; Aubrecht, Steinnocher & Huber, 2014; Smith, Martin & Cockings, 2016; Stathakis & Baltas, 2018) has been increasing the resolution of the temporal component and/or including a seasonal dimension for limited regional areas by mining geodata from different more or less conventional sources. Other authors have explored the contribution of big data such as mobile phone activity records (Deville *et al.*, 2014; Tatem *et al.*, 2014) or 'geotweets' (Patel *et al.*, 2017) for population mapping in selected countries, a task which is not without shortcomings and challenges. A relatively straightforward approach was proposed to estimate day- and night-time population distribution at high resolution for the cities in Urban Atlas (Freire, Florczyk & Ferri, 2015); yet its quality depends largely on that of ancillary datasets and availability of local parameters, and accuracy levels have not been assessed against independent sources.

The challenges posed by spatiotemporal mapping and modelling of population distribution cannot be addressed effectively by conventional data sources alone (e.g. official statistics and reference land use datasets). Significant advances in this field can only be attained if data from conventional data sources are combined with data from emerging, non-conventional data sources in a coherent methodological framework. Non-conventional data sources may include volunteered geographical information (Goodchild, 2007), web-based social networks (Aubrecht *et al.*, 2017), proprietary thematic databases, mobile phone operator data, or even navigation systems.

Several studies have documented the relevance and usefulness of mobile phone data in particular (see for example

Deville *et al.*, 2014; Steenbruggen, Tranos & Nijkamp, 2015), but one key problem remains data access, which is normally negotiated with the data providers by individual researchers for specific projects (Demunter & Seynaeve, 2017).

1.2 Scope and objectives

The ENACT project aims at developing and implementing a consistent and validated methodology to produce multi-temporal population density grid maps (or 'population grids', for short) for Europe. The final output of ENACT is a set of multi-temporal population grids that take into account the main seasonal and daily variations of population, consistent with the most recent censuses data (2011), and covering the whole of EU28. The target spatial resolution is 1 Km, which is sufficiently detailed for sub-regional and local scale analyses and applications. The target temporal resolution is day- and night-time for the 12 months of the years, hence resulting in a total of 24 population grids. These novel datasets are expected to not only expand the knowledge base of spatiotemporal population patterns across the continent but will be useful inputs to applications in various fields. These include assessment of human exposure to natural and technological hazards, assessment of demand for resources (e.g. energy and water), planning and modelling of transport, land use, regional economy and environment.

While the project is in its last phase, with the final outputs currently under validation, the purpose of this paper is to provide an overview of the ENACT project and its overall methodology. The final results of the project will be documented in a forthcoming publication.

2 Data and methods

The final outcome of the ENACT project is a set of 24 population grids, each representing night-time or daytime for each month of the year. This set of multi-temporal population grids are discrete and not representative of the whole daily or weekly cycles. Each grid represents a 'typical' working day of the month. The weekend variation is not addressed. The night-time slot represents an 'ideal' situation when everybody is assumed to be at home for rest/sleep, whereas the daytime slot refers to a situation when everybody is assumed to be at the location of their primary activity during core working hours. As such, intermediate daily variations of population are not taken into account (e.g. commuting, pre- or after-work activities, etc.). The reference year for population data is 2011, to ensure consistency with the latest round of censuses.

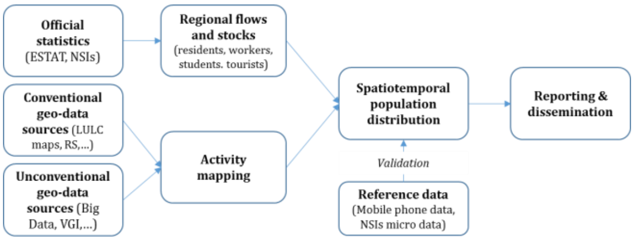
Although the working spatial resolution is 100 x 100 m, results are made available at 1 x 1 km due to difficulties in carrying a validation exercise at the working resolution. Releasing the data at the working resolution could mislead users regarding the actual precision and accuracy of the product.

The methodology is structured in four main tasks or phases. Table 1 describes these tasks briefly, and refers to their main inputs, outputs, and methods. Some tasks run in parallel (1 and 2), while others have dependencies (task 3 depends upon completion of tasks 1 and 2, and task 4 depends on completion of task 3). Figure 1 shows the overall workflow of the project and the next subsections provide further detail for each task.

Table 1: Main tasks of the ENACT project, brief description, key inputs, outputs and methods.

Task / phase	Description	Inputs	Outputs	Methods
1. Regional population flows and stocks	Estimation of flows and stocks of various population subgroups per region (NUTS3) and per month.	Official statistics (Eurostat, National Statistical Offices)	Set of 12 tables, one per each month of the year. Each tables has x records = n rows (regions) x k population subgroups.	Search and assembly of statistical data; tabular operations supported by spreadsheet and statistical packages.
2. Activity mapping	High-resolution (100m x 100m) mapping of location of socioeconomic activities (e.g. places of residence, employment, study, leisure).	Geodata from multiple sources (proprietary, free/open/public/voluntary repositories).	'ENACT base map', consisting of a refined version of the CORINE Land Cover Map 2012, with improved spatial and thematic detail; Set of complementary layers (hotel room density, density of shops, schools, etc.).	Search and assembly of geographical data; geodata pre-processing and fusion, supported by GIS.
3. Population disaggregation	Creation of population grids by disaggregating regional population stocks to grid level using location of activities as spatial proxies.	Outputs of tasks 1 and 2.	Two population grids for each month of the year (total of 24 population grids).	Spatial downscaling supported by GIS.
4. Validation	Comparison of ENACT's outputs with population grids derived from independent sources or methods.	Output of task 3; population grids derived from independent sources or methods (e.g. Mobile Phone Operator Data).	Goodness of fit measures.	Pre-processing and harmonization of reference population grids; Calculation of comparison indicators, supported by GIS and statistical packages.

Figure 1: Overall project workflow



2.1 Regional population flows and stocks

This task aims at constructing stocks of relevant population subgroups per month and per NUTS3 regions. We considered a total of 16 population subgroups, as follows:

- Residents;
- Employees per 11 economic sectors;
- Students per 2 categories;
- Non-active population;
- Inbound tourists.

Residents correspond to the number of registered residents within a region. This figure is obtained from Eurostat directly at NUTS3 level. *Employees* are broken down by 11 relatively broad economic sectors stemming from the NACE rev.2 classification of economic activities. The employment statistics were obtained from Eurostat and reflect the NUTS3 region of work. Gap filling was necessary to complete the available dataset from Eurostat. *Student* statistics reflect the region where students are enrolled in education institutions. Students are subdivided into 'below tertiary education' and 'tertiary education and above', but in both cases, numbers were only available per NUTS2 regions. Students below tertiary education were distributed among the respective NUTS3 regions based on the proportion of the relevant population age-groups. Higher education students were downscaled to NUTS3 regions based on the number of enrolled students per NUTS3 available from the European Tertiary Education Register (ETER). The *non-active* population subgroup refers to population not working nor studying, and comprehends retired population, children not attending kindergartens, unemployed, and inactive working-age population. The common denominator of this population subgroup is the likelihood that its members hang around residential areas for a significant share of the daytime. The estimation of this stock per NUTS3 involved the combination of data from various Eurostat tables.

Inbound tourists are defined broadly as visitors (thus temporary residents) in the region for any purpose, leisure and business altogether. Inbound tourists were derived by following a series of calculation steps. First, annual no. of nights-spent within a NUTS2 region (Eurostat) were disaggregated to NUTS3 regions based on the no. of bed-places per NUTS3 (Eurostat). Then, the NUTS3 annual no. of nights spent were broken down per month using regional (NUTS2 or NUTS3) seasonal curves constructed from data procured from every National Statistical Office (NSO). Finally, the average daily no. of inbound tourists is obtained by dividing the nights-spent per region and per month by the respective no. days in the month. For more details on the methods and results of the spatiotemporal mapping of tourists, see Batista e Silva *et al.* (2018).

Also from NSOs we obtained data allowing us to split inbound tourists in a country per country of origin. Tourists from countries outside the study area (i.e. EU28) represent added population to the existing stock. Tourists from the same country (domestic) or from countries within the study area (non-domestic) had to be subtracted from their regions of origin to avoid double counting of total population within the study area. In Figure 2 a visual representation of the flow of tourists between countries within the study area can be appreciated.

In sum, the seasonal variation of the total present population in a region is linked primarily to touristic flows, both inbound and outbound. Figure 3 shows, for one alpine region in Austria, the monthly variation in stock of present residents and inbound tourists resulting from the described procedure.

Figure 2: Origin-destination of non-domestic nights-spent.

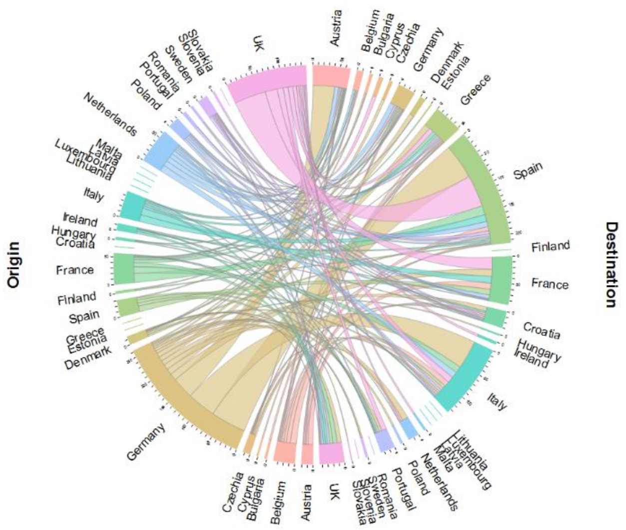
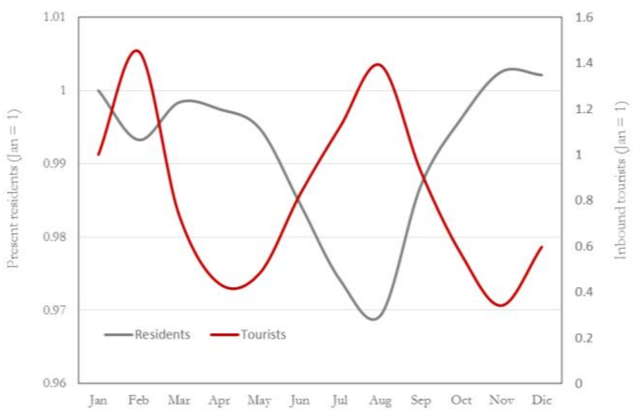


Figure 3: Monthly variation of the stock of present residents (axis on the left) and inbound tourists (axis on the right) in the region of Tiroler Unterland, Austria (NUTS3 code 'AT335'), indexed to the month of January.



2.2 Activity mapping

To map the locations of probable presence of population for the various sub-groups we resorted to multiple data sources, proprietary and open, conventional and non-conventional.

Land use and land cover (LULC) are widely used as ancillary variables (covariates) in fine resolution population mapping and modelling (Wardrop *et al.*, 2018). In the context of the ENACT project, we constructed our own 'base map' for the disaggregation of population stocks from regional to grid level. This base map is a refined version of the well-known CORINE Land Cover map, version 2012 (CLC 2012). The refinement of CLC 2012 includes two key components:

a) Spatial refinement;
b) Thematic refinement.

The *spatial refinement* consists of improving the spatial detail and geometric representation of the LULC classes. In practice, we reduced the minimum mapping unit from 25 hectares to 1-5 hectares (typically 1 ha for artificial surfaces, 5 hectares for all other LULC classes). The reduction was achieved by combining information from multiple geodata sources, and their integration with the original CLC 2012 in a sequential order. Input data have been selected and harvested upon compliance with the following criteria:

- Compatibility with CLC's LULC nomenclature (LULC class definitions);
- Reference year 2012 +/- 2;
- Higher spatial resolution than CLC 2012;
- Pan-European geographical coverage;
- Preferably free, open and documented data.

The input data sources are listed below:

- CLC products: CLC 2012 v 18.5, CLC Changes 2006-2012 and CLC Changes 2000-2006;
- Copernicus high resolution (HR) layers 2012: HR layer Forest type + Tree cover density, HR layer Permanent water bodies, HR layer Wetlands;
- TomTom Multinet 2014: Land Use layer + Built-up layer;
- JRC's European Settlement Map (ESM) (10m version, aggregated to 100 m reference grid);
- Urban Atlas 2012 (~ 700 Functional Urban Areas covered);
- OpenStreetMap (OSM) and TomTom Multinet 2014 as source of road network data.

The *thematic refinement* consists in increasing the breakdown of the LULC categories originally available in CLC. CLC uses a hierarchical nomenclature with 44 classes at level 3, however only 11 comprise artificial surfaces. The goal is to derive more specific subclasses of human activities, which can be linked to various population subgroups. A level 4 is therefore added to the CLC nomenclature.

Methods based on cartographic overlay between the original map and ancillary layers are applied to derived classes 11XX, 122X, 124X, 1421, and 1422. The breakdown of class 121 in 3 subclasses is based upon a machine learning classification approach. After the spatial refinement step, each 121 polygon was geometrically intersected with the road network from TomTom to obtain a finer set of polygons (~0.75 million). Each resulting polygon is characterized by a number of explanatory variables (weighted sum of Points of Interest from 16 categories) and other contextual variables related to population density, land use/land cover in the neighborhood, and distance to key transport features). The sources of geodata to construct these explanatory variables include:

- Proprietary geographical data: TomTom Multinet (transport and miscellaneous Points of Interest), PLATTS (energy), EuroRegionalMap from EuroGeographics (miscellaneous);
- Open and public sources: European Pollutant Release and Transfer Register (location of industries);
- Volunteered Geographical Data: OpenStreetMap (miscellaneous Points of Interest).

For about 1/3 of all the class 121 polygons we have determined their ground truth class based upon detailed national land use maps for a selection of countries and regions (COS for Portugal, SIOSE for Spain, COSW for Wallonie, and DUSAF for Lombardy), as well as land use polygons obtained from OpenStreetMap and TomTom. All this information was then used by Machine Learning classifiers to construct an explanatory model which was then be used to classify 121 polygons without ground-truth in the level 4 categories. An independent validation carried by comparing the automatic classification with human interpreted imagery yielded an overall accuracy of 74% with a Kappa of 0.53 (for

Figure 4: Differences between the original CLC 2012 (left) and the CLC after spatial and thematic refinement (right) for the area surrounding Brussels, Belgium.

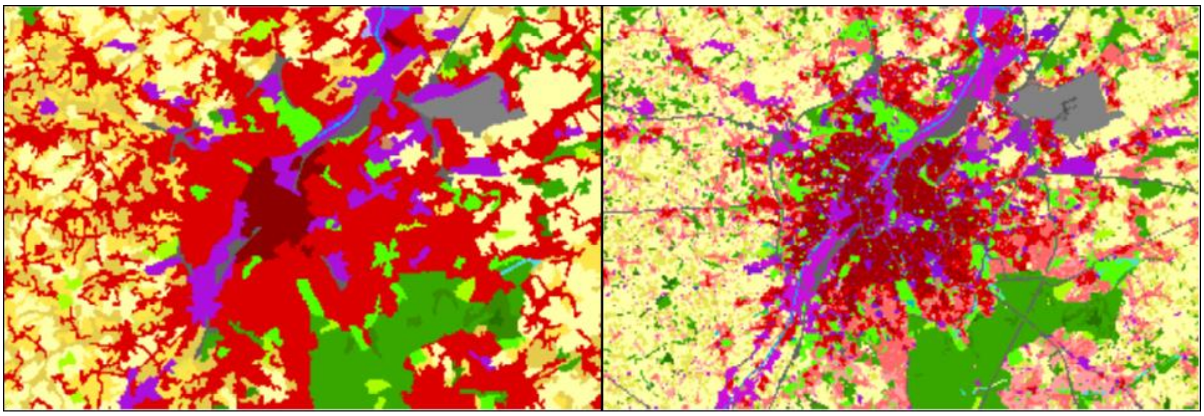
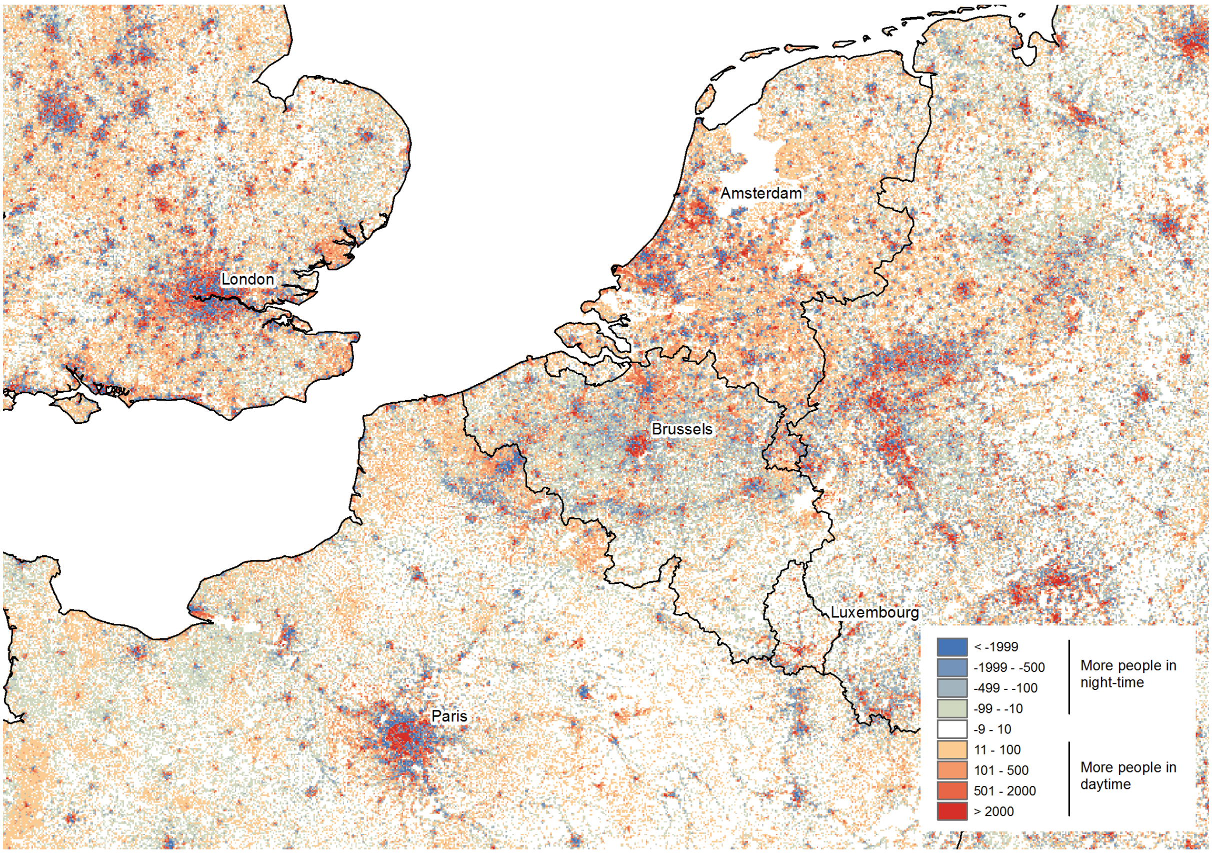


Figure 5: Difference in total population per 1 km^2 between day- and night-time (provisional results).



more details refer to Rosina *et al.* (2018)).

The 'activity mapping' task included as well the production of some complementary activity layers, such as those depicting touristic accommodation room density (based on data from online booking services), retail and food service density (constructed with data from TomTom Multinet Points of Interest), and location of schools. These additional layers were needed due to conceptual difficulties in integrating point-based data of hotels, shops, and schools in the polygon-based LULC map. These layers play a key role in the subsequent spatial allocation of population subgroups such as tourists, students, or shop workers.

2.3 Population disaggregation

Regional population stocks were disaggregated to grid level using the ENACT refined LULC map, and the complementary activity layers mentioned previously. Residents are assigned entirely to areas deemed residential, while the stock of tourists is assigned to locations of touristic accommodation. The sum of these 'gridded' residents and tourists constitutes the 'night-time' grid for each given month.

To produce daytime population grids, we allocate all population subgroups (excluding residents) to the relevant LULC classes and in some cases the allocation is also based on complementary activity layers. The spatial allocation, or disaggregation, is governed by a probability matrix establishing a link between each population subgroup and each LULC or activity class. The matrix was built based on expert judgement. The disaggregation is further guided by the local built-up densities as per the European Settlement Map.

Figure 5 shows the difference between day- and night-time grids at 1 km resolution, as per an early version of the results, showing marked spatial configurations.

2.4 Validation / cross-comparison

ENACT's population grids are being compared against data from at least two types of independent sources:

- Mobile phone usage density grids from mobile phone operators;
- Commuting data from which day and night-time population per municipality can be derived.

The comparison is being carried for areas for which data are available. Currently, we possess mobile phone operator data for Amsterdam (Jacobs-Crisioni *et al.*, 2014), Milan and Trentino (Barlacchi *et al.*, 2015), and Belgium (Demunter & Seynaeve, 2017). Commuting data is, instead, available for Italy, Portugal, Poland and Czech Republic.

The comparison is done on a case by case basis, as each data source has different degree of spatial, temporal and thematic resolutions. In most cases, the ENACT output will have to be aligned to match the definition and resolution of the independent source. For example, day and night-time population derived from commuting data does not incorporate tourists; or grids from mobile phone operator refer to usage density, rather than total population density.

Options to deal with these differences may include variable rescaling or careful design of comparison measures based on relative densities across time and space, rather than absolute ones. Moreover, ENACT's gridded values at 100 x 100 m will be aggregated to the spatial zoning system of the reference sources, and not the other way around, to avoid distortion of the reference data.

3 Early conclusions and way forward

Although the project is not finalized, some preliminary conclusions can already be drawn. Multi-temporal modelling of population distribution is an exceptionally data-intensive task, especially for large study areas. Data integration is challenging due to the required data volume and variety in terms of formats, definitions, nomenclatures and/or semantics.

The long and intricate workflow to combine such a variety of data inputs, each with its own inaccuracies, leads inevitably to a propagation and accumulation of error in the final product too. Knowing the accuracy of the produced datasets is necessary to inform the users of the product but also to determine the spatial scale at which it should be used. Therefore, designing a robust quality assessment strategy is no less important and challenging as the modelling per se. Unfortunately, a proper and systematic validation is difficult to implement due to the lack of truly comparable data. A set of cross-comparison metrics for selected areas, using independent datasets, is currently being laid to hopefully shed light on the reliability and plausibility of ENACT's outputs.

A planned follow-up includes the update of ENACT's grids to reflect population data from 2015. Additional work should target the improvement of the accuracy and temporal resolution, towards more continuous population grids. As it is likely that the quality and quantity of suitable input data will grow in future, it is also vital to attempt at developing a flexible modelling framework that could accommodate gradual data (and methodological) improvements.

Disclaimer

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

References

Ahola, T., Virrantaus, K., Krisp, J.M. & Hunter, G.J. (2007) A spatio-temporal population model to support risk assessment and damage analysis for decision-making. *International Journal of Geographical Information Science*. [Online] 21 (8), 935–953. Available from: doi:10.1080/13658810701349078.

Aubrecht, C., Aubrecht, D.Ö., Ungar, J., Freire, S., et al. (2017) VGDI – Advancing the Concept: Volunteered Geo- Dynamic Information and its Benefits for Population Dynamics Modeling. *Transactions in GIS*. [Online] 21 (2), 253–276. Available from: doi:10.1111/tgis.12203.

Aubrecht, C., Steinnocher, K. & Huber, H. (2014) DynaPop - Population distribution dynamics as basis for social impact evaluation in crisis management. *ISCRAM 2014 Conference Proceedings - 11th International Conference on Information Systems for Crisis Response and Management*. [Online] (May), 314–318. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-84905869676&partnerID=40&md5=c7c2013d66c885aedc410427b89c95da.

Barlacchi, G., De Nadai, M., Larcher, R., Casella, A., et al. (2015) A multi-source dataset of urban life in the city of Milan and the Province of Trentino. *Scientific Data*. [Online] 2, 150055. Available from: http://dx.doi.org/10.1038/sdata.2015.55.

Batista e Silva, F., Gallego, J. & Lavalle, C. (2013) A high-resolution population grid map for Europe. *Journal of Maps*. [Online] 9 (1), 16–28. Available from: doi:10.1080/17445647.2013.764830.

Batista e Silva, F., Marín Herrera, M.A., Rosina, K., Ribeiro Barranco, R., et al. (2018) Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*. [Online] 68, 101–115. Available from: doi:https://doi.org/10.1016/j.tourman.2018.02.020.

Bhaduri, B., Bright, E., Coleman, P. & Urban, M.L. (2007) LandScan USA: A high-resolution geospatial and temporal modeling approach for population distribution and dynamics. In: *GeoJournal*. [Online]. 2007 pp. 103–117. Available from: doi:10.1007/s10708-007-9105-9.

Demunter, C. & Seynaeve, G. (2017) Better quality of mobile phone data based statistics through the use of signalling information - the case of tourism statistics. In: *New Techniques and Technologies for Statistics 2017*. [Online]. 2017 Eurostat. p. Available from: doi:10.2901/EUROSTAT.C2017.001.

Deville, P., Linard, C., Martin, S., Gilbert, M., et al. (2014) Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences*. [Online] 111 (45), 15888–15893. Available from: doi:10.1073/pnas.1408439111.

Freire, S., Florczyk, A. & Ferri, S. (2015) Modeling Day-and Nighttime Population Exposure at High Resolution: Application to Volcanic Risk Assessment in Campi Flegrei. *Proceedings of the ISCRAM 2015 Conference*. [Online] (August). Available from: http://iscram2015.uia.no/wp-content/uploads/2015/05/5-2.pdf.

Gallego, F.J., Batista, F., Rocha, C. & Mubareka, S. (2011)

Disaggregating population density of the European Union with CORINE land cover. *International Journal of Geographical Information Science*. [Online] 25 (12), 2051–2069. Available from: doi:10.1080/13658816.2011.583653.

Goodchild, M.F. (2007) Citizens as sensors: The world of volunteered geography. *GeoJournal*. [Online]. 69 (4) pp.211–221. Available from: doi:10.1007/s10708-007-9111-y.

Jacobs-Crisioni, C., Rietveld, P., Koomen, E. & Tranos, E. (2014) Evaluating the impact of land-use density and mix on spatiotemporal urban activity patterns: An exploratory study using mobile phone data. *Environment and Planning A*. [Online] 46 (11), 2769–2785. Available from: doi:10.1068/a130309p.

Martin, D., Cockings, S. & Leung, S. (2015) Developing a Flexible Framework for Spatiotemporal Population Modeling. *Annals of the Association of American Geographers*. [Online] 105 (4), 754–772. Available from: doi:10.1080/00045608.2015.1022089.

Martin, D., Cockings, S. & Leung, S. (2010) Progress report: 24-hour gridded population models. In: *Paper presented at the European Forum for Geostatistics Conference 2010, Tallinn, Estonia, 5-7*. 2010 pp. 1–9.

McPherson, T.N. & Brown, M.J. (2004) Estimating daytime and nighttime population distributions in U.S. cities for emergency response activities. In: *Bulletin of the American Meteorological Society*. [Online]. 2004 pp. 557–566. Available from: doi:10.1215/9780822384625-001.

Patel, N.N., Stevens, F.R., Huang, Z., Gaughan, A.E., et al. (2017) Improving Large Area Population Mapping Using Geotweet Densities. *Transactions in GIS*. [Online] 21 (2), 317–331. Available from: doi:10.1111/tgis.12214.

Rosina, K., Batista e Silva, F., Vizcaino-Martinez, P., Marín-Herrera, M.A., et al. (2018) An improved European land use/cover map derived by data integration. In: *Proceedings of the AGILE Conference 2018*. 2018 Lund, Sweden. p.

Smith, A., Martin, D. & Cockings, S. (2016) Spatio-Temporal Population Modelling for Enhanced Assessment of Urban Exposure to Flood Risk. *Applied Spatial Analysis and Policy*. [Online] 9 (2), 145–163. Available from: doi:10.1007/s12061-014-9110-6.

Stathakis, D. & Baltas, P. (2018) Seasonal population estimates based on night-time lights. *Computers, Environment and Urban Systems*. [Online] 68, 133–141. Available from: doi:https://doi.org/10.1016/j.compenvurbsys.2017.12.001.

Steenbruggen, J., Tranos, E. & Nijkamp, P. (2015) Data from mobile phone operators: A tool for smarter cities? *Telecommunications Policy*. [Online] 39 (3–4), 335–346. Available from: doi:10.1016/j.telpol.2014.04.001.

Tatem, A.J., Huang, Z., Narib, C., Kumar, U., et al. (2014) Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria Journal*. [Online] 13 (1). Available from: doi:10.1186/1475-2875-13-52.

Wardrop, N.A., Jochem, W.C., Bird, T.J., Chamberlain, H.R., et al. (2018) Spatially disaggregated population estimates in the absence of national population and housing census data. *Proceedings of the National Academy of Sciences*. [Online] Available from: doi:10.1073/pnas.1715305115.