Mapineq

Mapineq Link: Environmental geospatial database module

Florianne Verkroost
University of Groningen

Giovanni Scotti Bentivoglio
University of Groningen

Bo Zhao
University of Oxford

Douglas Leasure
University of Oxford

Melinda C. Mills
University of Oxford / University of Groningen

Mapineq deliverable D1.2 September 2023

Funded by the European Union

Suggested citation:

Verkroost, F., Scotti Bentivoglio, G., Zhao, B., Leasure, D. & Mills, M.C. (2023). Mapineq Link: Environmental geospatial database module. Mapineq deliverables Mapineq deliverables. Turku: INVEST Research Flagship Centre / University of Turku. DOI: 10.5281/zenodo.10400861

Summary history		
Version	Date	Comments
1.0	19.09.2023	Draft submitted for review
2.0	27.09.2023	Final Revised Document
2.1	18.12.2023	DOI from Zenodo

Mapineq – Mapping inequalities through the life course– is a three-year project (2022-2025) that studies the trends and drivers of intergenerational, educational, labour market, and health inequalities over the life course during the last decades. The research is run by a consortium of eight partners: University of Turku, University of Groningen, National Distance Education University, WZB Berlin Social Science Center, Stockholm University, Tallinn University, Max Planck Gesellschaft (Population Europe), and University of Oxford

Website: www.Mapineq.eu

The Mapineq project has received funding from the European Union's Horizon Europe research and innovation programme under the grant agreement No. 101061645.

Acknowledgement:

This document was reviewed by Kadri Täht (Tallinn University) and Mijail Figueroa González (University of Turku) as part of Mapineq quality assurance procedures. The content of the document, including opinions expressed and any remaining errors, is the responsibility of the authors.

Publication information:

Executive summary

The Mapineq project focuses on the trends and drivers of intergenerational, educational, labour market and health inequalities across the life course. An aim of economic, social and health policy interventions is to diminish inequalities. To adequately study this, we require an understanding of how inequalities accumulate and are influenced by the natural environmental context at the local, regional and national level.

The Mapineq Link database is a rich open-access online interactive database containing information on geospatially linked structural and institutional policies. The database can be accessed through an API (Application Programming Interface) in order to link these data to the various research projects within and beyond Mapineq project.

Mapineq Link improves existing databases by: (1) harmonising and blending information from multiple data sources, (2) including geospatial indicators at the neighbourhood, city and/or regional level, (3) including geo-coded data to link to individual data; (4) providing better and broader subject coverage; (5) using objective information as opposed to subjective opinions from experts; (6) using diverse data sources from the bottom-up; (7) containing structural, socioeconomic, policy and environmental data; and, (8) providing an API for researchers, programmers, data scientists and web developers to actively and easily use the data. The current report is an extension of the Phase 1 of the Mapineq Link database (Leasure et al. 2023) that focussed on ten thematic modules focussing on structural and social and economic policy indicators (e.g., family, education, labour market, poverty).

In Phase 2 of the Mapineq Link project, we focus on environmental variables, which refers to the physical geospatial context where individuals live and work, which can include lighting, noise or pollution levels, traffic conditions or temperature. Our environmental database section contains **four data geo-spatial types**: (1) point, (2) line, (3) raster, and (4) polygon **across five thematic modules**: (1) earth and natural landscape, (2) built landscape, (3) administrative boundaries, (4) climate and environment; and (5) socio-economic indicators.

Mapineq Link will come with an online interactive dashboard and an API which allows: (1) researchers to query the database, (2) programmers and data scientists to retrieve data for specific geolocations or bespoke regions; (3) external web developers to easily integrate Mapineq data into their own interactive web applications; and (4) other users (e.g. the public) to view data and visualizations of a wide range of indicators. The Mapineq Link team is dedicated to FAIR principles to make our innovations Findable, Accessible, Interoperable, and Reusable.

Funded by
the European Union

Abbreviations

API	Application Programming Interface
Eurostat	Statistical Office of the European Union
FAIR	Findable, Accessible, Interoperable and Reusable
GADM	Globally-harmonised administrative boundaries
LAU	Local Administrative Units
NEET	Not in Education, Employment, or Training
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organisation for Economic Cooperation and Development
UN	United Nations

Content

Tables

Figures

Mapineq Link: Environmental geospatial database module

The Mapineq project focuses on the trends and drivers of intergenerational, educational, labour market and health inequalities across the life course, a core area of social, economic and policy research. A central aim of economic, social and health policy interventions is to diminish inequalities. In order to do so, we require an understanding of how inequalities may accumulate and how they are influenced by the natural environment (e.g., pollution, temperature) at the local, regional and national level. Environmental context can operate to aggravate or diminish inequalities. This report introduces Phase 2 of the Mapineq Link geospatial environmental database, a rich open-access database with API access that researchers can use to link a multitude of geo-linked policies to their own research projects.

Mapineq Link **improves existing databases** by: (1) harmonising and blending multiple sources, (2) including geospatial indicators at the neighbourhood, city or regional level, (3) including geo-coded data to link to individual data; (4) better and broader subject coverage; (5) objective information; (6) bottom-up use of diverse data sources; (7) structural, socioeconomic, policy and environmental data; and, (8) an API to link data.

Our **environmental database** section contains **four data geo-spatial types**: (1) point, (2) line, (3) raster, and (4) polygon **across five thematic modules**: (1) earth and natural landscape, (2) built landscape, (3) administrative boundaries, (4) climate and environment, and (5) socio-economic indicators.

Mapineq Link has an **API and online dashboard** which allows: (1) public and policy-makers, accessible material, (2) researchers to query the database, (3) programmers and data scientists to retrieve data for specific geolocations or bespoke regions; (4) external web developers to easily integrate Mapineq data into their own interactive web applications. The Mapineq Link team is **dedicated to FAIR principles** to make our innovations Findable, Accessible, Interoperable, and Reusable.

Funded by
the European Union

1. Introduction

The Mapineq project focuses on the trends and drivers of intergenerational, educational, labour market and health inequalities across the life course, a core area of social, economic and policy research. Life course domains include those spanning early childhood, parenting, childcare and family, transition to adulthood, education, labour market, occupational mobility to retirement. These life course events and domains are, however, experienced differently by individuals depending on the context they live or grew up in, which motivates the exploration of life course differences from a spatial perspective. While differences in experiences across the life course can relate to variation in individual and family traits, they are also shaped by local, regional, national and supranational opportunity structures and influenced by local spatial as well as cross-regional and cross-national variations in resources, institutions and norms across. As noted in the first module of the Mapineq database module (Leasure et al. 2023), the emergence and shaping of life course inequalities between individuals and groups will always be contingent on social and physical context, which motivates our pursuit of a rich spatial database on life course inequalities.

The current report is part of a series of reports that describe the different facets of the Mapineq Link geospatial database. It extends Phase 1, which examined social and economic policy institutional indicators (Leasure et al. 2023). The Mapineq Link geospatial database is a rich open-access database with API access that researchers can use to link a multitude of geo-linked policies to their own research projects. This report introduces Phase 2 of the database, focussing on environmental variables, which refers to the physical geospatial context where individuals live and work, which can include buildings, road networks, housing, noise or pollution levels, light or temperature. These environmental and policy layers will later be available and visualised through an open-access web interface planned in a later phase of the Mapineq project, released in 2024.

1.1. Research problem

For the purposes of this database, the **environment** is defined as a system encompassing all natural, biotic and abiotic and artificial (human-made) elements affecting any aspect of human life. The natural environment consists of living and non-living things provided by nature and occurring naturally, with the human environment referring to all forms and results of human intervention in the natural environment. Life course researchers have more prominently focussed on social and economic policies, such as the contextual database of the Mapineq Link module 1 (Leasure et al. 2023). Yet many aspects of the natural environment may be important in shaping inequalities, opportunities and multiple aspects of life events and trajectories. In this section we provide some illustrative examples of how researchers can study the relationship and impact of the natural and human environment and how their interaction fosters patterns of social, economic and health inequalities. Environmental context can be traced to obesogenic, polluted or urbanised environments or environments exposed to high hydrogeological risk. Among its several features, Mapineq Link, through its extensive and granular data on geospatial indicators, is designed to enable researchers to study and address the effects of the interplay of natural and human environments on inequalities over the life course.

As outlined in phase 1 of our project (Leasure et al. 2023), although some rich environmental databases already exist, they face several challenges that Mapineq Link aims to overcome by doing the following:

(1) **harmonising and blending multiple sources**, moving away from previous fragmented databases where information is spread across different locations and online databases, forcing researchers to access multiple unharmonized sources;

(2) **including fine-grained geospatial indicators** at the neighbourhood, city and/or regional level, incorporating more extensive variation and moving beyond previous databases that often include information at the unit of analysis of the national level only;

(3) **providing geo-coded data to link to individual data**, given that previous data have not often been geo-coded, meaning that individuals cannot be linked to their immediate ecosystem or distance from key predictors (e.g., distance to green space, school, pharmacy). As the database is developed, users who have access to geolocated individual-level data can use locations (e.g., latitude and longitude coordinates), to query our database to retrieve spatially-linked data for those locations across multiple data sources;

(4) **expanding coverage of topics**, given that many previous projects focussed on economic, family, housing and health but rarely combined all topics. We provide broader and deeper spectrum of detailed policies and ability to link to demographic and structural aspects disaggregated by key inequality stratifiers of age, sex, region, migrant status, education or labour market activity;

(5) including **objective information from multiple sources**, moving beyond previous valuable projects that drew on information from expert interviewers (e.g., Multilinks);

(6) adding **both top-down but also bottom-up diverse generated data** from population registries, surveys and other types of emerging data (e.g., environment, social media). Previous data has often relied on indicators derived from local, regional and national statistical offices or supranational organisations, thus missing the opportunity to generate indicators from the bottom-up data sources;

(7) **inclusion of structural, socioeconomic, policy and data on the natural and built-up environments.** In Phase 2 of the Mapineq Link project, we will include multiple new types of natural and built environment data, which is often not included in these types of databases.

(8) **developing an API (Application Programming Interface) in an interactive online dashboard** to allow researchers to query the online dashboard and database, programmers and data scientists to retrieve bespoke geospatial data; and, external web developers to integrate Mapineq Link data to their own web applications.

In the first instance our database will focus on the **European context and associated countries.** As data becomes available and the database evolves over time, we may expand to other regions of the world.

Funded by
the European Union

1.2. The role of environmental variables in Life Course research

Individuals are inherently embedded in a sum of natural and artificial elements that determine their development and influence their life course. Despite the awareness of the fundamental role of the environment in shaping individuals' socioeconomic and demographic outcomes, social science research rarely studies the interplay between specific environmental factors and such outcomes. An increased availability of environmental and geospatial data, and growing attention and concern with regard to environmental conditions, means that an expanding body of literature has started to investigate the impact of environmental indicators, measuring specific aspects of the physical geospatial context in which individuals live and work, on life course outcomes and domains. In this context, a crucial goal of this phase of the Mapineq Link project is to understand the role of environmental variables in life course domains in relation to educational, labour and health inequalities. To provide researchers with an indication of how this database might be applied, in this section we outline a brief selection of the countless ways through which natural and artificial environments can influence inequalities, providing some insights, backed by the existing literature, on the actual role of diverse environmental conditions in life course events.

1.2.1. Air quality and educational achievement

A large body of evidence documents the negative direct or indirect effects of toxic air and air pollution on educational achievements and cognitive performance during childhood. Direct effects have been studied to examine the impact of air quality on several dimensions of cognitive abilities and development (Cserbik et al. 2020, Alvarez-Pedrerol et al. 2017, Kweon et al. 2018). Others have examined how air pollution can affect educational achievement indirectly through a decrease in school attendance or sleep quality (Barone-Adesi et al. 2015).

The Mapineq study of Bernardi and Conte Keivabu (2023), available on Mapineq's website, investigates the role of socioeconomic status (SES) in the impact of poor air quality on educational inequalities among Italian secondary-education students. In particular, it assesses the presence of a differential exposure to poor air quality by parental SES and notes heterogeneity by parental SES in the effect of air quality on school results.

1.2.2. Climate change and migration

Extreme weather events including flooding, wildfires and high temperatures means that climate change is a growing topic across multiple life course domains. There is a vast empirical literature on the effect of climate shocks and extreme environmental conditions on migration. The decision to move in response to climate stress is not trivial (Cattaneo et al. 2019) and needs to be framed in the context of interrelated economic and cultural factors, taking into account the characteristics of the household and the life course stage (Van Praag 2021). In a conceptual framework, the household decision to migrate depends on the impact of weather anomalies on resources, both material and intangible, and exposure to shocks (Kaczan and Orgill-Meyer 2019). The interplay between these aspects can affect the form of migration (in terms of destination and timing after the event), also considering the different effects of fast onset and slow onset climatic events (United Nations Framework Convention on Climate Change, 2012).

1.2.3. Climate change, heat waves and mortality

Heat waves have always represented a public-health hazard, but as their frequency is increasing due to climate change, understanding in-depth the effects on health and mortality has crucial policy implications. The definition of heat wave varies in the literature, but it generally takes into account both the intensity and the duration of the event (Gasparrini and Armstrong 2011). While many studies have shown that heat waves disproportionately affect the mortality of vulnerable demographic groups, such as the elderly (Astrom et al. 2011), evidence on which socioeconomic group most exposed to extreme heat is mixed and varies across studies and is often related to a complex interplay of factors related to housing and other related conditions (e.g. Michelozzi et al. 2005, Xu et al. 2013). For instance, air pollution plays a confounding role in the effect of heat waves on mortality, and not including it in the analysis can lead to a significant bias in the results (Analitis et al. 2014).

1.2.4. Light and noise pollution and sleep

Human generated environmental light and noise pollution can have extremely detrimental impacts on sleep and subsequently daytime functioning and health (Halperin 2014; Dumont& Beaulieu 2007). Given that light is the primary environmental cue for the biological clock to regulate the body's day and night cycle, inappropriate light exposure has been shown as clinically relevant and related to mood, sleep and circadian disorders (Dumont & Beaulieu 2007). Noise, particularly caused by nocturnal road traffic and proximity to road, has been measured as a significant cause of sleep disorders and health (Pirrera, De Valck & Cluydts 2010). Poor sleep provokes measurable biological changes in the form of stress response of endocrine and metabolic perturbations and been related to cardiometabolic, psychiatric and negative social and educational outcomes in children and adults (Halperin 2014). The result of light and noise pollution is daytime tiredness, annoyance and mood changes, decreased well-being and cognitive performance which can be linked to negative family and social interactions as well as adverse health.

1.2.5. Housing prices and fertility patterns

Housing prices are an important human environment economic and spatial indicator of the liveability of an environment, and as such may affect life course events. A growing strand of empirical literature addresses the effect of fluctuations in housing prices on individual and couple's fertility choices and preferences. The evidence shows opposite effects for home-owners and renters, thus revealing different mechanisms underlying fertility choices (Clark and Ferrer 2019). For the former group, an increase in house-prices increases in turn home equity value, leading to a positive effect on birth rates, whereas for the latter group a house-price increase has a negative (Dettling and Kearney 2014) or not significant (Lovenheim and Mumford 2013) effect on fertility. Although similar results emerge from different settings, the effect is clearly mediated by the socioeconomic conditions of the household.

These topics represent only a fraction of potential applications of the Mapineq Link environmental data, with the expectation that it will lead to many more diverse and novel studies.

2. Mapineq Link Geospatial Database

As noted in our previous description of module 1 (Leasure et al. 2023), the Mapineq Link geospatial database goes beyond being a data download portal by providing a cloud-based toolbox for spatial queries of the innovative combinations of data that will ultimately be included in the database. This provides capability for place-based spatial linkages among traditional surveys, registries, and census-based data, and also spatial linkages with non-traditional digital and remote sensing datasets that will be added to the database in later phases of the project. Mapineq Link also provides functionality for cloud-based calculation of zonal statistics within specific geographies (e.g., NUTS regions), and distances to points-of-interest such as schools or clinics. For Phase 1 of database development, we focused on building and deploying the database infrastructure in the cloud and populating the database with socio-economic policy indicators and measures at national and subnational levels.

For the current Phase 2 of database development, we focused on populating the database with environmental variables and measures at national and subnational levels as well. In addition to the socio-economic policy indicators of Phase 1, these variables test as well as showcase the ability and advantages of Mapineq when it comes to its geospatial elements. In addition, anticipating the eventual launch of Mapineq Link to the public, we focused on the design and development of the APIs that will give researchers and developers access to this database quickly and easily. Within the five thematic modules, we prioritised so-called case study data sets of four types of geospatial data: 1) point data; 2) line data; 3) raster data; and 4) polygon data. As Mapineq Link develops further, environmental indicators of each of these four types will be added to the database on a continuous basis. The case study data sets have been collected in a number of ways, depending on the data source – some using an API and others using (bulk) downloading links. The Mapineq database reuses these formats for data and metadata with some modifications for harmonisation, documentation and spatial linkage purposes. For every dataset included in Mapineq Link, we also include the data license and citation from the data provider. We comply with data license terms in how we obtain, use, and redistribute data, and we provide license terms to end-users of Mapineq Link for every dataset so that they are aware of any license restrictions for their own uses of the data.

2.1. Geographic Units of Analysis

When we focus on policies, the geographical unit of analysis matters not only in terms of the implementation of those policies but also the data collection regarding their efficacy. In terms of the former, policies can be implemented and relevant at the national, regional or local level. In terms of the latter, data on the policies themselves or on outcomes influenced by them can be collected only at the national but also at regional and fine-grained geo-coded levels. For instance, federal government spending, national educational policies (e.g. required years of schooling), or national labour market policies (e.g. minimum wages), tend to be measured at the national level. If, however, regional or local authorities allocate spending or educational policies are handled by region or province/state, the relevant geographic unit of variables might change. Furthermore, many variables that are influenced by national-level policies, such as the unemployment rate or country of birth of residents may be more relevant to measure at a fine-grained lower geographic unit. The Mapineq Link geospatial database accommodates data queries across these geographic levels and subnational variations. Furthermore, it facilitates various widely used administrative boundaries datasets such as European NUTS regions, International Territorial Levels used by OECD and the United Kingdom, as well as GADM (globally-harmonised administrative boundaries). What is essential is the harmonisation of subnational information not only between different definitions of administrative regions, but also within these definitions (e.g., when it comes to inconsistencies over time). For Phase 1, we focused on NUTS regions and OECD International Territorial Levels. For the current Phase 2, GADM boundaries have also been added to Mapineq Link. While this database will mainly focus on Europe, we added GADM boundaries to more easily facilitate comparisons with countries outside of Europe, which is often important and desirable from a substantive research point of view. As new data becomes available, it will remain a living database, where we will continue to include additional geographic unit frameworks as database development progresses.

2.1.1. NUTS Divisions

The Nomenclature of Territorial Units for Statistics (NUTS) is a geocode standard, which was adopted in 2003 by the European Union. The Mapineq Link database includes the NUTS polygon geometries and geocodes all EuroStat data to these boundaries. These boundaries can also be used to produce zonal statistics summarising other spatial layers that will be incorporated into the database in later phases (e.g. phase II incorporates environmental spatial layers).

A hierarchy of three NUTS levels (i.e. NUTS1, NUTS2 and NUTS3) was established by Eurostat. The subdivisions do mostly, but not necessarily, correspond to administrative divisions within the countries. Below the three NUTS levels are local administrative units (LAUs). The current NUTS classification, valid from 1 January 2021, lists 92 regions at NUTS 1, 242 regions at NUTS 2, 1166 regions at NUTS 3 level, and 99,151 local administrative units (LAUs). The NUTS framework provides spatial units of analysis that allow for a comparative spatial setup of European countries and that captures existing administrative boundaries. The wide usage of this system in EU, national and regional data collection makes this the most prevalent spatial classification system for data collected thus far.

2.1.2. OECD International Territorial Levels

The territorial levels used by the OECD roughly correspond to NUTS regions, but the spatial extent includes OECD countries outside of the European Union. There are also differences in Germany, United Kingdom, France, Greece, The Netherlands, Belgium, and Iceland where the hierarchy of TL levels do not map directly to NUTS levels. The polygon geometries for OECD territorial levels are not publicly available, so we have a request pending through project collaborators at OECD to access these boundaries for use in the Mapineq database. Currently, the database maps OECD TL geocodes to spatial geometries of NUTS regions. For non-European countries where NUTS regions are not available, the Mapineq database may need to link TL region codes to GADM boundaries.

2.1.3. GADM Database of Global Administrative Areas

GADM administrative boundaries (gadm.org) are openly available for free with global coverage and general alignment with NUTS regions in the European Union. The Mapineq database may need to rely on GADM boundaries for any countries included that are outside of the European Union, including some OECD countries. While Mapineq focuses on the EUR, it is important to have spatial geometries defining the boundaries of non-European OECD geographies so that these national and subnational indicators can be spatially linked to other datasets in the database and to facilitate cross-country comparisons outside of Europe. The descriptions of the types of data are illustrative rather than exhaustive and does not include all variables listed (see Appendix A and Dashboard for full details).

3. Environmental Variables

The environmental module of the database consists of variables across five substantive themes and four geo-spatial data types.

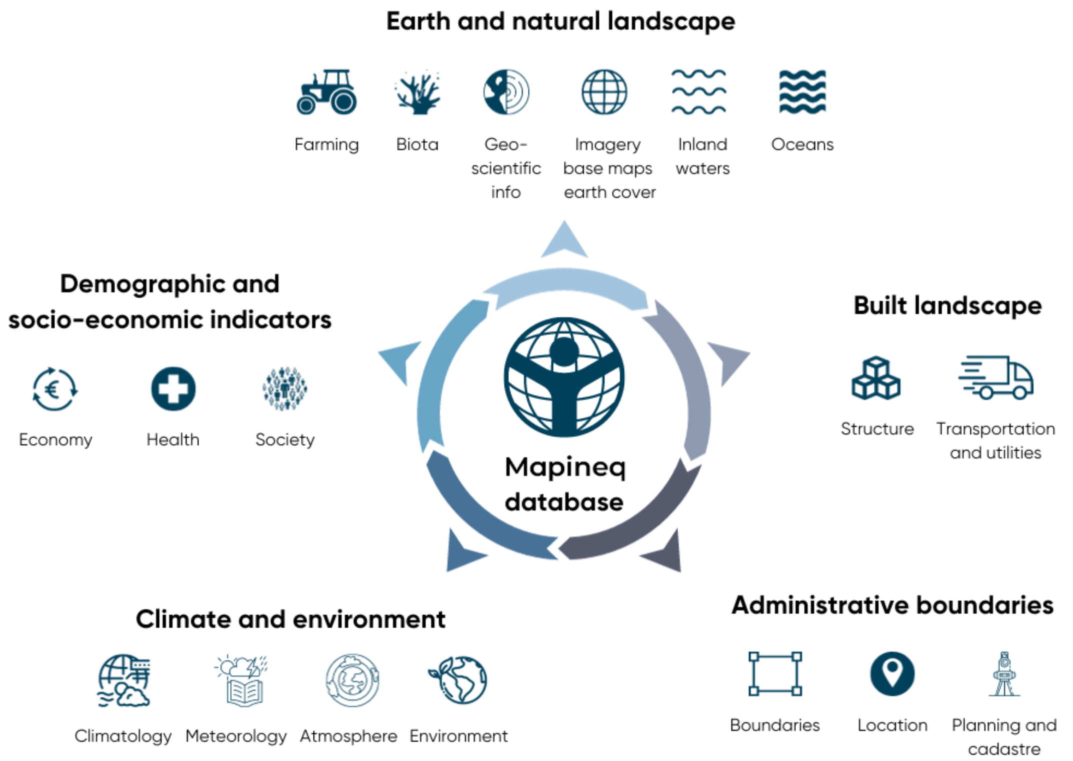
These **five substantive themes** aggregate parts of the ISO19115 Topic Category[1] (see Figure 1) and include:

1. **earth and natural landscape** (including farming, biota, elevation, geo-scientific information, imagery base maps earth cover, inland waters and oceans);
2. **built landscape** (including structure, transportation and utilities);
3. **administrative boundaries** (including boundaries, location, planning and cadastres);
4. **climate and environment** (including climatology, meteorology, atmosphere and environment);
5. **socio-economic indicators** (including economy, health and society).

[1] https://apps.usgs.gov/thesaurus/thesaurus-full.php?thcode=15

Funded by
the European Union

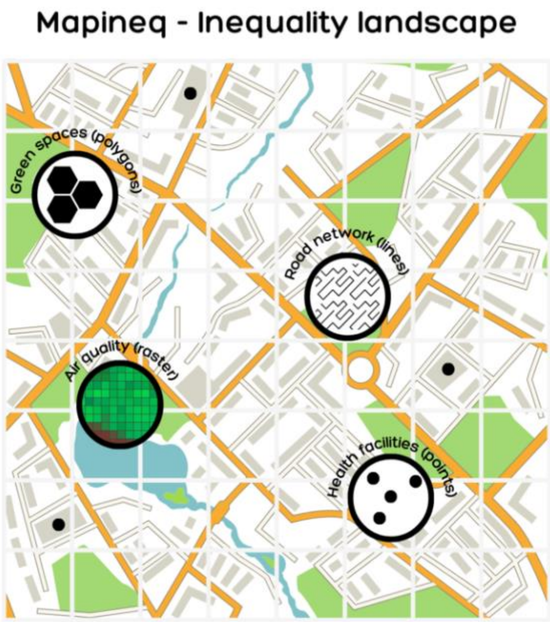
Figure 1. Visualisation of the five thematic modules of Module 2 environmental indicators



The **four geo-spatial data types** include: (1) point data; (2) line data; (3) raster data; and (4) polygon data. These are represented visually in Figure 2.

Given that a certain thematic module is not always linked to one particular geo-spatial data type and vice versa, we structure this section by the four geo-spatial data types while linking back to the thematic modules. The geo-spatial data types are our starting point to be able to develop the functionality of the Mapineq database to link with different types of geospatial data and traditional indicators, which is the basis for future extensions of the thematic modules as Mapineq Link develops further.

Figure 2. Visualisation of the four geospatial data types of Module 2 environmental indicators



3.1. Point data

For Phase 2 of Mapineq Link, we focused on a number of point type of geospatial data. Point data are often found in the second thematic module concerning the built landscape (e.g., when it comes to facilities and buildings) or the fourth thematic module concerning climate data (i.e., when the data come from specific point stations).

Examples of point data currently included in Mapineq Link are levels of ozone as well as UV radiation, for different periods of time as well as measuring devices and profiles. That is, we include total column ozone as well as vertical ozone profiles (lidar, ozone sonde, rocket sonde and Umkehr-N14), and for UV radiation we include broadband, multiband and spectral data. These data are available from the world Ozone and Ultraviolet Radiation Data Centre (WOUDC).[2] The geospatial points refer to measurement stations worldwide.

While ozone and UV radiation are merely starting examples of point data that will be included in Mapineq Link, they more generally reflect a substantive relevance of such data. Exposure to particular ozone or UV radiation levels may be related to other climate indicators on a macro-level, such as the DNA of animals and plants as well as the earth's surface (Norval et al., 2011; Roy, 2017). On a micro-level, these types of environmental exposures might be related to outcomes in other domains of life, mainly the human health domain. For example, ozone exposure has previously been linked to problems in the cardiovascular and respiratory systems, mortality and new-onset asthma and small-airway function in children (Nuvolone et al., 2018; Sousa et al., 2013) as well as premature ageing of the lungs (Lippmann, 1989). Similarly, while UV radiation may benefit human health by stimulating the vitamin D synthesis, it can also damage human health through

[2] https://woudc.org/data/explore.php?lang=en

photosensitivity and cutaneous malignancies (MacKie, 2000) and eye disease (Gallagher and Lee, 2006).

Other examples of point data that we are considering for future inclusion in the database are locations of schools, hospitals, clinics, and daycare facilities. Accessibility to services depends on their locations relative to the populations they serve, and geographic disparities in access can underpin inequalities in health outcomes (Weiss et al., 2020). We are also considering including locations of grocery stores and fast-food restaurants to facilitate research into the geographic distributions of obesogenic environments as a driver of health inequalities (Giskes et al., 2011; Kirk et al. 2010). Inclusion of these data depends on our ability to retrieve them with consistent coverage from across Europe, and we are currently assessing the feasibility of this.

Finally, also looking ahead to Phase 3 of Mapineq Link, we are planning to include housing prices from cadastral and/or real estate websites as point data into the database. As aforementioned, the availability and pricing of housing is related to a variety of inequalities in other domains of life, such as fertility (Clark and Ferrer 2019), gentrification (Wilhelmsson et al., 2022) and commuting (Mayock, 2016) – which is why these data would be important to incorporate.

3.2. Line data

For Phase 2 of Mapineq Link, we focused on a number of line type of geospatial data. Line data are often found in data contained in the second thematic module, namely the built landscape and particularly infrastructure and transportation.

An example of line data currently contained in the database is a detailed network of the road network in Great Britain based on the Ordnance Survey.[3] In the future, these data could be extended to include European countries (which were not [yet] available without openly [i.e., without login credentials] at the time of writing). The geospatial lines refer to the network of roads classified by the National or Local Highway authority, and are updated bi-yearly.

Road networks data are a critical component to support calculation of "distance to" measures within the database. For example, a user could provide point locations and ask to retrieve the distances to the nearest health care facilities or other points-of-interest that are in the database. For many research questions, Euclidean distances (i.e., straight line distance) would be less suitable than road distances (i.e., driving distance). Our aim is to support user queries to calculate both Euclidean distances and driving distances from user-provided point locations to various points-of-interest in the Mapineq Link database.

Similarly, the road data included in the Mapineq Link currently provide a mere example of line data. Other types of line data we are considering to include in the future are rivers, trails and streets. These data are interesting from a substantive point of view because road proximity and connectivity are related to various life course domains and outcomes (Boothe and Shendell, 2008), such as term low birth weight (Dadvand et al., 2014); reduced occurrence of animal species such as birds (Brotons and Herrando, 2001); heavy metal occurrence in soil and consequently plant health (Khalid et al., 2018); and

[3] https://www.ordnancesurvey.co.uk/products/os-open-roads

Funded by the European Union

neurologic disease incidence (Yuchi, 2020). Beyond these physical consequences, road proximity and connectivity might also influence subjective (inequalities in) well-being, for example when it comes to job satisfaction and commuting time (Handy and Thigpen, 2019) or access to everyday facilities (Jones et al., 2016) and urban centres (Weiss et al., 2018).

3.3. Raster data

For Phase 2 of Mapineq Link, we focused on a number of raster type of geospatial data. Raster data are usually linked to the fourth module of climate and environment data, when weather station data are interpolated across a pre-specified grid of smaller areas. Raster data currently included in the database are precipitation, solar radiation, water vapour pressure, wind speed, and temperature (minimum, average and maximum).[4] These data are obtained from WorldClim and available for 1970-2000. Version 2.1. was released in January 2020, and monthly data are collected at the 10 minute (~340 squared kilometres) resolution. The geospatial rasters refer to these blocks of earth surface.

Raster data are usually found as remotely sensed and topographic data, including satellite and aerial imagery. Such data are substantively relevant, especially when it comes to mapping (the consequences of) climate change throughout the past decades. For example, while the effects on migration have been discussed earlier, climate and environment may also affect inequalities in health and well-being in different ways – for example when it comes to green space (Markevych, 2017), electrification (Falchetta et al., 2020), infrastructure (Zhou et al., 2022) and development more generally (Puttanapong et al., 2022).

Satellite-based measures of buildings are another important raster dataset that we will incorporate into the database. Two great examples are the Global Human Settlement layer from the European Commission (Melchiorri et al., 2018) and the World Settlement Footprint from the German Aerospace Centre (Esch et al., 2022). These datasets quantify the coverage of buildings–and sometimes building heights–for small grid cells globally and their changes through time. These space-based observations of the built environment are fundamental for understanding the inequality landscape in terms of geographic distributions of building densities, urbanization, settlement sizes, proximity to undeveloped areas, and so forth. A top priority set of raster datasets that we will include in the database is gridded population data. WorldPop is one widely-used example that provides annual population estimates (until 2020) disaggregated by age and sex for every 100 m grid square globally (Tatem 2017). High-resolution gridded population data provide the ability to aggregate 100 m grids to generate population estimates into spatial units appropriate for a given research question. This could range from administrative areas (e.g., NUTS regions) to school or hospital catchment areas that are provided by users. This provides an important source of population denominators for producing per capita inequality indicators for relevant subnational areas.

[4] https://www.worldclim.org/data/worldclim21.html

Funded by the European Union

3.4. Polygon data

For Phase 2 of Mapineq Link, we focused on a number of polygon types of geospatial data. Polygon data are usually found in data that refer to bounded areas, such as the natural boundaries (e.g., wetlands, forests) of the first thematic module as well as the administrative boundaries (e.g., postal codes, NUTS regions) of the third thematic module – which usually also are the geographical units of the socio-economic and demographic data of the fifth theme.

The first set of example data currently included in Mapineq Link are collected from Natural Earth, funded by the North American Cartographic Information Society (NACIS), and contain information about roads, railroads, airports, ports and urban areas at the 1:10M scale.[5] The second set of data are collected from the Global Lakes and Wetlands Database, provided by the WWF and the Center for Environmental Systems Research (University of Kassel, Germany).[6] The first level (GLWD-1) consists of the world's largest lakes and reservoirs, whereas the second level (GLWD-2) consists of permanent open water bodies not included in GLWD-1. There are 3,067 lakes of an area larger than 50 squared kilometres as well as 654 reservoirs of storage capacity over 0.5 cubic kilometres in GLWD-1; GLWD-2 contains open water bodies that are at least 0.1 squared kilometres in surface area. The geospatial polygons refer to the shapes of these lakes and wetlands.

Most often, polygon data refer to bounded areas such as the aforementioned lakes and wetlands, but also for example forests and cities. These data could be relevant to explore a variety of questions in the socio-economic inequality domain, such as when it comes to the relationship with urban forestation (Mills et al., 2016) or street tree abundance (Lin et al., 2006).

When it comes to further extensions in the future, Ookla internet speed tests data are a top priority for integration into the Mapineq database. These data provide quarterly aggregated results for millions of internet speed tests conducted by the public. These results are provided within 600 m grids for users connected via mobile networks as well as fixed "landline" networks. Geographic differences in access to high-speed broadband and mobile networks is a contributing factor for inequalities in job opportunities, access to information to improve both education and health (Lucendo-Monedero, 2019). This often manifests as an urban/rural divide, and it has been shown to be a barrier to compliance with public health guidance for the COVID-19 pandemic (Stern et al., 2019; Chiou and Tucker 2020).

3.5. Future plans

The Mapineq Link database has now been populated with examples of all data formats that it supports including geospatial points, lines, and polygons, as well as the raster data described above. It also contains tabular data from OECD and EuroStat that are linked to NUTS and GADM administrative units. This allows the Mapineq team to begin development work on the various types of geospatial queries and data linkages that the database will ultimately support. Throughout the development process, we will continue to integrate

[5] https://www.naturalearthdata.com/downloads/

[6] https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database

Funded by
the European Union

additional data sets into the database to broaden the scope of research topics that Mapineq Link will be able to support.

The following types of spatial queries and data linkages will be developed in the upcoming period to enable the Mapineq user community to:

— Select NUTS or GADM administrative units to return OECD and/or EuroStat indicators for these regions;

— Select NUTS or GADM regions to summarize environmental data within those regions (e.g. sum, average, length, count);

— Provide point coordinate(s) to extract OECD and/or EuroStat indicators as well as environmental data from those locations or within a user-defined radius around the locations;

— Provide point coordinate(s) to calculate "distance to" measures relative to the nearest features-of-interest (i.e. from the environmental data sets) based on Euclidean (straight-line) or road distances;

— Provide polygons (e.g. user-defined catchment areas) to summarize environmental data within each polygon;

— Provide polygons (e.g. user-defined catchment areas) to retrieve NUTS or GADM regions that intersect or fall completely within each polygon;

These tools are the foundation that will elevate the Mapineq Link database above a simple data download portal or web mapping portal. We are striving to give users the power of spatial data for their own research locations without the burden of mastering skills in spatial data science or the need to scour the internet for suitable geospatial data from many disparate sources.

Next we will focus on developing the Mapineq application programming interface (API) to deliver these tools to end users. This mode of access will be suitable for web developers, data scientists, and statistical programmers who want to bulk query the database programmatically. This is a prerequisite to begin work on the Mapineq dashboard which will use the API to query the database.

Finally, we will develop a web mapping data dashboard that will provide easy access to these data and tools for the general public. Users of the dashboard will be able to explore the database visually on an interactive map, explore relationships among variables, and click to create points and polygons that can be used to query the database as described above.

4. Conclusions

This report introduced Phase 2 of the second module of the ambitious Mapineq Link database project. Our data collection phase offers multiple contributions to the existing research landscape by providing a 'one stop shop' where researchers, policy makers, programmers, data scientists and web developers can use and extract data under FAIR principles. Our improvements to existing comparable databases include harmonisation and combination of multiple data sources, more fine-grained geospatial indicators; an ability to link geospatial data to individual data; broader subject coverage; bottom-up (i.e., inputting data from surveys, population registers) and top-down (i.e., national and supranational) use of diverse data sources.

Phase 1 focused on preparing the database infrastructure and collecting socio-economic policy indicators. In this Phase 2, we extended the database infrastructure and additionally collected primary environmental data of 4 data geo-spatial types: (1) point, (2) line, (3) raster, and (4) polygon across 5 thematic modules: (1) earth and natural landscape, (2) built landscape, (3) administrative boundaries, (4) climate and environment, and (5) socio-economic indicators. The inclusion of these data in Mapineq Link makes an important contribution because humans' environments, whether social or physical, natural or built, may influence their decisions and outcomes in such a myriad of ways that they may contribute to the accumulation of socio-economic inequalities throughout the life course.

Given that Mapineq Link is still under development, we note that there are various limitations and ongoing extensions of this work. First, the environmental data shown and described in this report serve as merely examples of the expansive data types and themes that are available and will be integrated into the database. In this phase, we prioritised the actual development and testing of the database, and will continue to include environmental indicators from a substantive point of view as Mapineq Link develops further. Second, our data collections currently showcase a cross-section of data available at the time of collection. However, ideally we would consider also the temporal continuum and collect data at various points in time for longitudinal purposes. Third, while Mapineq Link currently only contains socio-economic policy indicators (Phase 1) and environmental indicators (Phase 2), we will focus on extending the database to include innovative socio-economic indicators from commercial and unconventional sources (e.g. housing and rental prices) in Phase 3. Fourth, the next phase will also be where we implement the functionality of the API to explore and search the data in an accessible and user-friendly manner.

Funded by
the European Union

References

Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., Suades-González, E., Donaire-Gonzalez, D., Cirach, M., & Sunyer, J. (2017). Impact of commuting exposure to traffic-related air pollution on cognitive development in children walking to school. Environmental pollution, 231, 837-844.

Analitis, A., Michelozzi, P., D'Ippoliti, D., De'Donato, F., Menne, B., Matthies, F., & Katsouyanni, K. (2014). Effects of heat waves on mortality: effect modification and confounding by air pollutants. Epidemiology, 15-22.

Åström, D. O., Bertil, F., & Joacim, R. (2011). Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. Maturitas, 69(2), 99-105.

Barone-Adesi, F., Dent, J. E., Dajnak, D., Beevers, S., Anderson, H. R., Kelly, F. J., & Whincup, P. H. (2015). Long-term exposure to primary traffic pollutants and lung function in children: cross-sectional study and meta-analysis. PloS one, 10(11), e0142565.

Bernardi, F., & Keivabu, R. C. (2023). Poor air at school and educational inequalities by family socioeconomic status. Max Planck Institute for Demographic Research.

Boothe, V. L., & Shendell, D. G. (2008). Potential health effects associated with residential proximity to freeways and primary roads: review of scientific literature, 1999–2006. Journal of Environmental Health, 70(8), 33-41.

Brotons, L., & Herrando, S. (2001). Reduced bird occurrence in pine forest fragments associated with road proximity in a Mediterranean agricultural area. Landscape and urban planning, 57(2), 77-89.

Cattaneo, C., Beine, M., Fröhlich, C. J., Kniveton, D., Martinez-Zarzoso, I., Mastrorillo, M., & Schraven, B. (2019). Human migration in the era of climate change. Review of Environmental Economics and Policy.

Chiou, L. and Tucker, C. (2020) Social Distancing, Internet Access and Inequality. Working Paper 26982. National Bureau of Economic Research. Available at: https://doi.org/10.3386/w26982.

Clark, J., & Ferrer, A. (2019). The effect of house prices on fertility: Evidence from Canada. Economics, 13(1), 20190038.

Cserbik, D., Chen, J. C., McConnell, R., Berhane, K., Sowell, E. R., Schwartz, J., & Herting, M. M. (2020). Fine particulate matter exposure during childhood relates to hemispheric-specific differences in brain structure. Environment international, 143, 105933.

Dadvand, P., Ostro, B., Figueras, F., Foraster, M., Basagaña, X., Valentín, A., & Nieuwenhuijsen, M. J. (2014). Residential proximity to major roads and term low birth weight: the roles of air pollution, heat, noise, and road-adjacent trees. Epidemiology, 518-525.

Dettling, L. J., & Kearney, M. S. (2014). House prices and birth rates: The impact of the real estate market on the decision to have a baby. Journal of Public Economics, 110, 82-100.

Dumont, M. and C. Beaulieu. (2007). Light exposure in the natural environment: relevance to mood and sleep disorders, Sleep Medicine, 8: 557-65.

Esch, Thomas, Elisabeth Brzoska, Stefan Dech, Benjamin Leutner, Daniela Palacios-Lopez, Annekatrin Metz-Marconcini, Mattia Marconcini, Achim Roth, and Julian

Zeidler. (2022). World Settlement Footprint 3D - A First Three-Dimensional Survey of the Global Building Stock. Remote Sensing of Environment, 270: 112877.

Falchetta, G., Pachauri, S., Byers, E., Danylo, O., & Parkinson, S. C. (2020). Satellite observations reveal inequalities in the progress and effectiveness of recent electrification in sub-Saharan Africa. One Earth, 2(4), 364-379.

Gallagher, R. P., & Lee, T. K. (2006). Adverse effects of ultraviolet radiation: a brief review. Progress in biophysics and molecular biology, 92(1), 119-131.

Gasparrini, A., & Armstrong, B. (2011). The impact of heat waves on mortality. Epidemiology (Cambridge, Mass.), 22(1), 68.

Giskes, K., van Lenthe, F., Avendano-Pabon, M. and Brug, J. (2011), A systematic review of environmental factors and obesogenic dietary intakes among adults: are we getting closer to understanding obesogenic environments? Obesity Reviews, 12: e95-e106.

Halperin, D. (2014). Environmental noise and sleep disturbances: A threat to health? Sleep Science, 7: 209-12.

Handy, S., & Thigpen, C. (2019). Commute quality and its implications for commute satisfaction: Exploring the role of mode, location, and other factors. Travel Behaviour and Society, 16, 241-248.

Jones, T., Horton, D., Mullen, C., Pooley, C., Scheldeman, G., Strano, E., & Tight, M. (2016). The role of street network connectivity and access to everyday facilities in shaping everyday walking and cycling in English cities. In Strategies for Sustainable Mobilities (pp. 15-38). Routledge.

Kaczan, D. J., & Orgill-Meyer, J. (2020). The impact of climate change on migration: a synthesis of recent empirical insights. Climatic Change, 158(3-4), 281-300.

Khalid, N., Hussain, M., Young, H. S., Boyce, B., Aqeel, M., & Noman, A. (2018). Effects of road proximity on heavy metal concentrations in soils and common roadside plants in Southern California. Environmental Science and Pollution Research, 25, 35257-35265.

Kirk, S.F.L., Penney, T.L. and McHugh, T.-.-L.F. (2010), Characterizing the obesogenic environment: the state of the evidence with directions for future research. Obesity Reviews, 11: 109-117.

Kweon, B. S., Mohai, P., Lee, S., & Sametshaw, A. M. (2018). Proximity of public schools to major highways and industrial facilities, and students' school performance and health hazards. Environment and Planning B: Urban Analytics and City Science, 45(2), 312-329.

Leasure, D., M.C. Mills, S. Wagner & F. Verkroost. (2023). Mapineq Link: Institutional policy and structural geospatial database. Mapineq deliverables. Turku: INVEST Research Flagship Centre / University of Turku. Available at: https://Mapineq.eu/wpcontent/uploads/2023/07/D1.1_MapineqLinkPhase1_v2.pdf

Lin, J., Wang, Q., & Li, X. (2021). Socioeconomic and spatial inequalities of street tree abundance, species diversity, and size structure in New York City. Landscape and Urban Planning, 206, 103992.

Lippmann, M. (1989). Health effects of ozone a critical review. Japca, 39(5), 672-695.

Lovenheim, M. F., & Mumford, K. J. (2013). Do family wealth shocks affect fertility choices? Evidence from the housing market. Review of Economics and Statistics, 95(2), 464-475.

Lucendo-Monedero, A.L., Ruiz-Rodríguez, F. and González-Relaño, R. (2019). ‘Measuring the digital divide at regional level. A spatial analysis of the inequalities in digital development of households and individuals in Europe’, Telematics and Informatics, 41, pp. 197–217.

MacKie, R. M. (2000). Effects of ultraviolet radiation on human health. Radiation Protection Dosimetry, 91(1-3), 15-18.

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., & Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. Environmental research, 158, 301-317.

Mayock, T. (2016). Wages, housing prices and commutes. Real Estate Economics, 44(1), 258-300.

Melchiorri M, Florczyk AJ, Freire S, Schiavina M, Pesaresi M, Kemper T. (2018). Unveiling 25 Years of Planetary Urbanization with Remote Sensing: Perspectives from the Global Human Settlement Layer. Remote Sensing, 10(5):768.

Michelozzi, P., De Donato, F., Bisanti, L., Russo, A., Cadum, E., DeMaria, M., & Perucci, C. A. (2005). The impact of the summer 2003 heat waves on mortality in four Italian cities. Eurosurveillance, 10(7), 11-12.

Mills, J. R., Cunningham, P., & Donovan, G. H. (2016). Urban forests and social inequality in the Pacific Northwest. Urban forestry & urban greening, 16, 188-196.

Norval, M., Lucas, R. M., Cullen, A. P., De Gruijl, F. R., Longstreth, J., Takizawa, Y., & Van Der Leun, J. C. (2011). The human health effects of ozone depletion and interactions with climate change. Photochemical & Photobiological Sciences, 10(2), 199-225.

Nuvolone, D., Petri, D., & Voller, F. (2018). The effects of ozone on human health. Environmental Science and Pollution Research, 25, 8074-8088.

Pirrera, S., De Valck, E. and R. Cluydts. (2010). Nocturnal road traffic noise: A review on its assessment and consequences on sleep and health, Environment International, 36(5): 492-498.

Puttanapong, N., Luenam, A., & Jongwattanakul, P. (2022). Spatial analysis of inequality in Thailand: Applications of satellite data and spatial statistics/econometrics. Sustainability, 14(7), 3946.

Roy, S. (2017). Impact of UV radiation on genome stability and human health. Ultraviolet light in human health, diseases and environment, 207-219.

Sousa, S. I. V., Alvim-Ferraz, M. C. M., & Martins, F. G. (2013). Health effects of ozone focusing on childhood asthma: what is now known–a review from an epidemiological point of view. Chemosphere, 90(7), 2051-2058.

Stern, M.J., Adams, A.E. and Elsasser, S. (2009). Digital Inequality and Place: The Effects of Technological Diffusion on Internet Proficiency and Usage across Rural, Suburban, and Urban Counties. Sociological Inquiry, 79: 391-417.

Tatem, A.J. (2017). WorldPop, open data for spatial demography, Scientific data, 4.

UNFCCC (2012) Slow onset events Technical Paper

Funded by
the European Union

Van Praag, L. (2021). Can I move or can I stay? Applying a life course perspective on immobility when facing gradual environmental changes in Morocco. Climate Risk Management, 31, 100274.

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., & Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature, 553(7688), 333-336.

Wilhelmsson, M., Ismail, M., & Warsame, A. (2022). Gentrification effects on housing prices in neighbouring areas. International Journal of Housing Markets and Analysis, 15(4), 910-929.

Xu, Y., Dadvand, P., Barrera-Gómez, J., Sartini, C., Marí-Dell'Olmo, M., Borrell, C., & Basagaña, X. (2013). Differences on the effect of heat waves on mortality by sociodemographic and urban landscape characteristics. J Epidemiol Community Health, 67(6), 519-525.

Yuchi, W., Sbihi, H., Davies, H., Tamburic, L., & Brauer, M. (2020). Road proximity, air pollution, noise, green space and neurologic disease incidence: a population-based cohort study. Environmental Health, 19(1), 1-15.

Zhou, Y., Li, X., Chen, W., Meng, L., Wu, Q., Gong, P., & Seto, K. C. (2022). Satellite mapping of urban built-up heights reveals extreme infrastructure gaps and inequalities in the Global South. Proceedings of the National Academy of Sciences, 119(46), e2214813119.

Appendix A. List of measures included in Mapineq Link version 2.0

Table 1. Overview of environmental geospatial data included currently in Mapineq Link Phase 2

Theme	Source	Indicator	Description	Geotype	License
Earth and natural landscape					
	WWF and the Center for Environmental Systems Research, University of Kassel, Germany	glwd_1	Level 1 (GLWD-1) comprises the 3067 largest lakes (area ≥ 50 km2) and 654 largest reservoirs (storage capacity ≥ 0.5 km3) worldwide, and includes extensive attribute data.	Polygon	https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database
		glwd_2	Level 2 (GLWD-2) comprises permanent open water bodies with a surface area ≥ 0.1 km2 excluding the water bodies contained in GLWD-1. The approximately 250,000 polygons of GLWD-2 are attributed as lakes, reservoirs and rivers.		
Built landscape					
	Ordnance Survey	oproad_gb	OS Open Roads is a high-level view of the road network, from motorways to country lanes in Great Britain. The links represent an approximate central alignment of the road carriageway and include roads classified by the National or Local Highway authority.	Line	https://www.ordnancesurvey.co.uk/products/os-open-roads
	NACIS	natural_earth_airports	1:10M cultural vector of world-wide airports in the public domain (derived from Mile High Club)	Point	https://www.naturalearthdata.com/about/terms-of-use/
		natural_earth_roads	1:10M cultural vector of world-wide roads in the public domain (derived from CEC North America Environmental Atlas)	Line	
		natural_earth_railroads	1:10M cultural vector of world-wide railroads in the public domain (derived from CEC North America Environmental Atlas)	Line	

Funded by
the European Union

		natural_earth_ports	1:10M cultural vector of world-wide ports in the public domain (derived from High Seas)	Point	
		natural_earth_urban_areas	1:10M cultural vector of urban areas in the public domain (derived from MODIS)	Polygon	
	Eurostat	ENV_WASFAC	Number and capacity of recovery and disposal facilities by NUTS 2 regions	Administrative regions (to map as polygons)	https://ec.europa.eu/eurostat/about-us/policies/copyright
Administrative boundaries					
	NUTS	nuts0	National boundaries	Polygon	https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units
		nuts1	Major socio-economic regions		
		nuts2	Basic regions for the application of regional policies		
		nuts3	Small regions for specific diagnoses		
	GADM	gadm_410		Polygon	https://gadm.org/license.html
Climate and environment					
	WorldClim	wc2.1_10m_prec_x	Precipitation (mm) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)	Raster	https://www.worldclim.org/about.html
		wc2.1_10m_vapr_x	Water vapor pressure (kPa) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		
		wc2.1_10m_wind_x	Wind speed (m/s) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		

		wc2.1_10m_tmin_x	Minimum temperature (degrees Celcius) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		
		wc2.1_10m_tmax_x	Maximum temperature (degrees Celcius) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		
		wc2.1_10m_tavg_x	Average temperature (degrees Celcius) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		
		wc2.1_10m_srad_x	Solar radiation (kJ/m2/day) in the month of January (x = 01) to December (x = 12) at 10 minutes resolution (~340 km2)		
	WOUDC	TotalOzone	Total column ozone (column ozone)	Point	https://www.canada.ca/en/environment-climate-change/corporate/transparency.html
		UmkehrN14	Vertical ozone profile (umkehr N-value)		
		OzoneSonde	Vertical ozone profile (ozone sonde)		
		RocketSonde	Vertical ozone profile (rocket sonde)		
		Spectral	UV radiance (spectral)		
		Lidar	Vertical ozone profile (lidar)		
		Broad-band	UV radiance (broadband)		
		Multi-band	UV radiance (multiband)		
	Eurostat	CLI_IAD_TD	Mean near surface temperature deviation (source: EEA)	Administrative regions	https://ec.europa.eu/eurostat/about-us/policies/copyright

Funded by
the European Union

				(to map as polygons)	
		ENV_AC_AEINT_R2	Air emissions intensities by NACE Rev. 2 activity		
Socio-economic indicators					
	OECD	C5_1	General government expenditure on environment protection, per person, USD 2015 PPP	Administrative regions (to map as polygons)	https://www.oecd.org/termsandconditions/
	Eurostat	CEI_CIE011	Persons employed in circular economy sectors	Administrative regions (to map as polygons)	https://ec.europa.eu/eurostat/about-us/policies/copyright
		CEI_CIE012	Private investment and gross added value related to circular economy sectors		
		CEI_CIE020	Patents related to recycling and secondary raw materials		
		CEI_GSR010	Consumption footprint		
		CEI_GSR011	Greenhouse gases emissions from production activities		
		CEI_GSR020	EU self-sufficiency for raw materials		
		CEI_GSR030	Material import dependency		
		CEI_PC020	Material footprint		
		CEI_PC030	Resource productivity		

Funded by
the European Union

		CEI_PC031	Generation of municipal waste per capita		
		CEI_PC032	Generation of waste excluding major mineral wastes per GDP unit		
		CEI_PC034	Waste generation per capita		
		CEI_PC035	Food waste		
		CEI_PC040	Generation of packaging waste per capita		
		CEI_PC050	Generation of plastic packaging waste per capita		
		CEI_SRM010	Contribution of recycled materials to raw materials demand - end-of-life recycling input rates (EOL-RIR)		
		CEI_SRM020	Trade in recyclable raw materials		
		CEI_SRM030	Circular material use rate		
		CEI_WM010	Recycling rate of all waste excluding major mineral waste		
		CEI_WM011	Recycling rate of municipal waste		
		CEI_WM020	Recycling rate of packaging waste by type of packaging		
		CEI_WM060	Recycling rate of waste of electrical and electronic equipment (WEEE) separately collected		
		CLI_ACT_NOEC	Share of zero emission vehicles in newly registered passenger cars (source: EAFO, DG MOVE)		

Funded by
the European Union

		CLI_IAD_LOSS	Climate related economic losses by type of event (source: EEA)		
		ENV_AC_CEPSGC	Intermediate consumption of environmental protection services by institutional sector and NACE Rev. 2 activity		
		ENV_AC_CEPSGH	Final consumption expenditure on environmental protection services by institutional sector		
		ENV_AC_EPIGG	Environmental protection investments of general government by environmental protection activity		
		ENV_AC_EPNEIS	National expenditure on environmental protection by institutional sector		
		ENV_AC_EPTRF	Environmental protection transfers by environmental protection activity and institutional sector		
		ENV_AC_EXP1R2	Environmental protection expenditure - million euro and million units of national currency		
		ENV_AC_EXP2	Environmental protection expenditure - euro per inhabitant and % of GDP		
		ENV_AC_EXP3	Environmental protection expenditure - % of pollution prevention, % of gross fixed capital formation, % of output		
		ENV_AC_PEPSGG	Production of environmental protection services of general government by economic characteristics		
		ENV_AC_TAX	Environmental tax revenues		
		ENV_AC_TAXENER	Energy taxes by paying sector		

Funded by
the European Union

		ENV_AC_TAXIND2	Environmental taxes by economic activity (NACE Rev. 2)		
		ENV_ESST_GG	Environmental subsidies and similar transfers from general government, by environmental activity, sector of recipient and ESA category of transfer		
		ENV_ESST_GGCP	Environmental subsidies and similar transfers from general government to corporations, by environmental activity, ESA category of transfer and NACE Rev. 2 activity of recipient		
		ENV_ESST_RW	Environmental subsidies and similar transfers from the Rest of the World to the domestic economy, by environmental activity and ESA category of transfer		
		ENV_WASPACR	Recycling rates of packaging waste for monitoring compliance with policy targets, by type of packaging		
		TEN00135	National expenditure on environmental protection		
		TEN00136	Environmental protection investments of total economy		
		TEN00139	Energy taxes		
		TEN00141	Environmental tax revenues		

Funded by
the European Union