Content

Tuesday,

15 November 2016

KEYNOTE

Data integration for mapping and monitoring the SDGs

Linus Bengtsson (Flowminder, Karolinska Institutet)

SESSION 1: UN-DSG AND UN-GGIM

Progress on the Global Statistical Geospatial Framework

Martin Brady (Australian Bureau of Statistics, Australia)

In the first week of August 2016, the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) adopted the 5 principles of the Global Statistical Geospatial Framework (GSGF). This is an important milestone for the UN-GGIM and for the statistical and geospatial communities globally.

The Global Statistical Geospatial Framework (GSGF) was proposed to the UN-GGIM by the UN Expert Group on the Integration of Statistical and Geospatial. The proposal followed several years of global collaboration and discussion by the statistical and geospatial communities and a round of formal global consultation across both National Statistical and National Mapping agencies. The global consultation received an overwhelmingly positive response, and the endorsement at the UN-GGIM had an unprecedented number of interventions, all supporting the work on the framework.

The 5 key principles in the framework are designed to facilitate the integration of statistical and geospatial information by ensuring consistent geospatial enablement of statistical and administrative data using common fundamental infrastructure. These principles provide the foundation of geospatial enablement of statistical and administrative data. The framework will be essential in supporting the 2020 round of Population Censuses, the Sustainable Development goal indicators, and provide the bridge to data more traditionally in the domain of the geospatial community, such as data on the natural and built environment.

This talk will review the recent progress to obtaining endorsement and look to the future opportunities and work associated with the framework.

Progress of UN-GGIM: Europe Working Group A on Core Data

François Chirié, Dominique Laurent (IGN, France)

UN-GGIM: Europe Working Group A has mandate to deal with core data specifications and quality, production issues, funding and data availability. Core data can be seen as the authoritative, harmonized and homogeneous framework data which both national and international users need to either fulfil their requirements or to geo-reference and locate their own thematic data.

The first step was to select core data themes among the 34 INSPIRE themes. This selection was mainly based on a user requirement investigation, focussing on the Sustainable Development Goals (SDG), and was completed by a comparison with core data themes selected by other initiatives. This user requirement survey was carried out by means of bibliographic research, mainly on the descriptions of INSPIRE use cases, and of expert interviews. In practice, core data should be the most necessary, most common, priority data required to analyse, monitor and achieve the SDGs, either directly or indirectly. Assessing and deciding

these priorities was achieved through setting up "use case maps" for each INSPIRE theme. After long discussions, WG A decided to select as core data themes: Addresses, Administrative units, Cadastral Parcels, Geographical Names, Hydrography, Transport Network, Elevation, Land Cover, Ortho-imagery, Statistical units, Buildings, Land Use, Area Management, Utility and Governmental Services.

The next step is to define data product specification for the 14 themes that have been selected, using INSPIRE data specification as starting point. In general, the work will consist to extract core content from INSPIRE (by selecting the most useful feature types and attributes in INSPIRE models or by clarifying or restricting the INSPIRE scope) and to add quality requirements in order to get more homogeneous content in Europe.

Europe Work Group B on Data Integration – Support of the better integration of geospatial information and statistics and the UN SDG monitoring

Pier-Giorgio Zaccheddu (Federal Agency for Cartography and Geodesy (BKG), Germany)

The aim of UN-GGIM: Europe is to ensure that the National Mapping And Cadastral Authorities (NMCAs) and National Statistical Institutes (NSIs) in the European UN Member States, the European Commission and its associated agencies work together to contribute to the more effective geospatial management and its integration with other information. The Executive Committee of UN-GGIM: Europe believes that there is scope for the regional entity to focus on how geospatial data can enhance sustainable development and the 2030 Agenda in Europe.

UN-GGIM: Europe Work Group B on Data Integration (WG B) is chaired by Germany and deals with the integration of geospatial data (including cadastral parcels) with other information. The focus of WG B in the past year has included a review of current European Interoperability Frameworks and best practice guidance for interactions between international organizations. Furthermore, an investigation of any side-effects induced by data combinations has been conducted. The status of the work of the WG according to the work plan 2015 – 2018 will be presented. The WG envisages to complete all tasks in November 2016.

The new Work Plan 2017 – 2020 for UN-GGIM:Europe, including new tasks for WG on Data Integration, was approved by the Plenary on 5 October 2016. It is focused on the support of the "geospatial dimension" of the UN Sustainable Development Goals (UN SDGs). UN-GGIM: Europe will work closely with the European Member States which are nominated as experts to Working Group on Geospatial Information of the Inter-Agency and Expert Group on SDG (IAEG-SDG). As a European contribution to the global process on developing a framework for monitoring indicators, UN-GGIM: Europe will – through the UN-GGIM: Europe WG on "Data Integration" – ensure a two-way interaction with the IAEG-SDG Working Group on Geospatial Information.

On the one hand the WG on Data Integration will ensure that the IAEG-SDG WG has access to existing work and ongoing working mechanisms in Europe related to achieving the goals and monitoring indicators. On the other hand, the WG on "Data Integration" will develop geospatial methodologies and approaches on monitoring, based on the European INSPIRE implementing rules and its technical specifications and on the SDG indicators from the IAEG-SDG WG, making it available to the European authorities responsible for monitoring. It will be necessary to elaborate how INSPIRE data and services can be used for the UN SDG monitoring.

Monitoring Agenda 2030 through a geospatial lens

Marie Haldorson (Statistics Sweden, Sweden)

In September 2015 the 2030 Agenda for Sustainable Development was adopted with an overarching principle that no one should be left behind. To support implementation at all levels, the 2030 Agenda included the need to exploit the contribution to be made by a wide range of data, including Earth observations and geospatial information.

In March 2015 at its forty-sixth session, the United Nations Statistical Commission created an Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs), which is composed of representatives from a regionally-balanced group of Member States and includes regional and international agencies as well as other key stakeholders, such as civil society, academia and the private sector, as observers. The IAEG-SDGs was tasked with providing a proposal for a global indicator framework (and associated global and universal indicators) for the follow up and review of the 2030 Agenda to be considered by the Statistical Commission. At the forty-seventh session of the Commission, in March 2016, the Global indicator framework was agreed upon by Member States.

Although the development of the Global indicator framework has primarily been based on a statistical data input-output approach, the need for 'geographic location' in a new era of data needs is well recognized. Many national statistical offices now understand that geospatial information, Earth observations and other Big Data are able to provide new and consistent data sources and methodologies to integrate multiple 'location-based' variables to support and inform official statistics and the indicators for the SDGs. Geography and location provides an important link to enable a richer picture of our countries, and what is happening in and across them. It enables data from diverse sources to be brought together to unleash their combined power in analysis and decision making.

To meet the ambitions and demands of the 2030 Agenda, it is necessary for the Global indicator framework to adequately and systematically address the issues of alternative data sources and methodologies, including geospatial information and Earth observations in the context of geographic location. Thus, at its forty-seventh session in March 2016, the IAEG-SDGs noted that the integration of geospatial information and statistical data will be key for the production of a number of the indicators. As a means to address these issues the creation of a Working Group on Geospatial Information, reporting to the IAEG-SDGs, is required.

The presentation will give an overview of the objectives and tasks of this new working group.

SESSION 2: COOPERATION BETWEEN NMAs AND NSIs

Population and Housing Statistics - The use of 3D models

Vilni Verner Holst Bloch (Statistics Norway, Norway)

Location, location, location. These are the three arguments used in housing prices, access to green areas, exposure to noise, solar influx, local climate, air pollution, flooding, and a wide range of other areas. To make an analysis of these topics one need a georeferenced point based register with the statistical units of interest, a 3D model of the geographic area of interest, and models to make indicators. This study explores the possibilities in applied georeferenced point based statistics and future cooperation between national statistical institutes (NSIs) and national mapping agencies (NMAs). Geographic 3D data is getting more and more accessible, cheaper and with higher resolutions and more possibilities of making statistics relevant down to smaller and smaller geographic areas. The use of detailed terrain data may contribute to and fill out gaps in the search for search for sustainable development goals for the next decades on a local to global level.

A questionnaire for collecting opinions for Statistical Units Data Specifications (UNGGIM-Europe, WGA)

Ignacio Duque (National Statistical Institute, Spain)

During the meeting in Paris in January 2016 the "Work Group A" (UN-GGIM-Europe) took a decision about the Core Data Themes, including Statistical Units.

Core data can be seen as the authoritative, harmonized and homogeneous framework data which both national and international users need to either fulfil their requirements or to geo-reference and locate their own thematic data. In the latter case, core data may be used as a framework on which other richer, more detailed, thematic geospatial and statistical data would rely. UN-GGIM: Europe believes that core data should be produced once for national and regional uses with maximum resolution, and would then be provided to international users if necessary through generalizing and aggregating processes. Core data is the data that is the most widely used, either directly or as a framework. Due to the United Nations context, WG A decided to focus the user requirements survey on the SDG-related requirements. Core data will be the data of main value for the UN Sustainable Development Goals (SDGs).

'Statistical Units' are the geographic part of a wide range of statistical data. The geography of statistical units is the mandatory bridge that connects the territory and statistical data. Simply combined with basic population information, theme SU provides a location associated with a number of persons and possibly with their characteristics (gender, age, etc.). Statistical Units may also be combined with more specialized statistics, such as socio-economic data or human health data, allowing various analyses about poverty, employment, education, health etc.

This theme is starting point for almost all studies at medium or large scales (small denominator scales), enabling deciders to identify the areas with major issues on a given topic. As Administrative units, NUTS/LAU may be used for money allocation. The INSPIRE theme 'Population Distribution' (PD), though of wide use for SDGs, has not been considered as core data because it is not geographic information. Theme PD is statistical information that may be combined with geographic data, generally themes 'Administrative Units' or 'Statistical Units' in order to perform powerful analysis.

The idea is to collect comments and suggestions for this document of specifications with a very brief questionnaire (with background documents). The presentations explain the questionnaire: aim, detailed information, target respondents and time span for collecting the forms. The questionnaires was designed after a round of consultation with Coordinators of Group A, Eurostat, EFGS steering committee and others.

Collaboration between French NMA and NSI (from NMA point of view)

François Chirié, Dominique Laurent (IGN, France)

IGN France (French NMA) has many relationships with INSEE (French NSI). First, IGN and INSEE have common interest in statistics: in 2012, IGN merged with the IFN (French Forest Inventory) and so, is now in charge of the public task to collect, manage and publish the statistical data of the forest inventory.

IGN carries out tasks requested by INSEE; for instance, IGN redefined urban units, according to United Nations recommendations, using both topographic data from IGN (mainly buildings) and population data from INSEE.

IGN and INSEE are also partners in the upgrade of statistical units data, called IRIS. IRIS are enumeration districts of 2 000 inhabitants. Purpose of this upgrade is to provide more accurate geometry to these IRIS by realigning their boundaries on the IGN background large scale topographic data.

IGN's mission is not only to produce data but it is also to integrate data from various origins. IGN has therefore an ambitious collaborative strategy and is developing partnerships in order to update and enrich its datasets. Some partnership with INSEE already took place e.g. IGN has used INSEE data about public facilities in order to populate a collaborative platform on "establishments open to the public". And there is room for further partnership as INSEE has a database named RIL containing information about addresses and buildings that could be a very relevant data source to enrich the IGN topographic data base and the national address database (BAN).

Resource Centre for GIS and Geospatial Information in Statistics Norway

Trine Haagensen (Statistics Norway, Norway)

In order to produce cost-effective land use statistics covering Norway, existing cartographic databases and registers has been used. Statistics Norway has created a hierarchical classification system, "Standard for classification of land use and land cover", which is based on both national and international standards and nomenclatures. The method applied is based on utilising the highest quality data sources available, but where no optimal data source exists, the next best quality data sources are used. In practical terms, the method is an automatic geographic information system (GIS) that defines, classifies and assembles the data into a hierarchy. In the statistics on land use and land resources all areas with buildings are classified as built-up. In addition, the building types within a built-up area determine the classification of that area.

SESSION 3: INTEGRATION OF GEOGRAPHY IN THE STATISTICAL PRODUCTION PROCESS

Geostat 2: Building and maintaining a point-based geospatial reference framework for statistics – preliminary conclusions from the project

Jerker Moström (Statistics Sweden, Sweden)

GEOSTAT 2 is a two year ESSnet grant project launched in February 2015. The aim of the project is to foster a better integration of statistics and geospatial information in order for the statistical community to provide more qualified descriptions of society, economy and environment in a spatial context.

One of its main objectives is to propose a generic model for a national point-based geospatial reference framework for statistics, building on national address, buildings and/or dwelling registers, suitable for statistics in the widest possible sense.

But what is a point-based geospatial reference framework and what are the challenges and benefits of using it? Is it even possible to propose a generic model considering the great variety of practices among statistical institutions throughout Europe?

The presentation will give a brief overview of the preliminary findings of the project so far. It will outline the concept of a point-based foundation for statistics and some of the operational aspects regarding setup and maintenance of such an infrastructure. Finally, the presentation will also briefly outline how the results from the project should be understood in relation to the Global Statistical Geospatial Framework.

Geostat 2: Some experiences and conclusions – applying GSBPM to geospatial statistical data production

Rina Tammisto, Marja Tammilehto-Luode, Tuuli Pihlajamaa (Statistics Finland, Finland)

A Generic Statistical Business Process Model (GSBPM) has been introduced in statistical agencies for describing statistical production processes. In the Geostat 2 project, the GSBPM was applied to the geospatial statistical production process. A coarse level model was created to understand the GSBPM and the relations between ordinary statistics and geospatial statistics production processes.

In order to fit the GSBPM to cover geospatial production processes, quite a lot of interpretations of current guidelines are required. However, Statistics Finland found the interpretations possible. Even on a very detailed level, our interpretation covered the whole production related to certain single geospatial statistics. The challenge was, however, to transmit one interpretation to a generic interpretation, which could even be utilised by another organisation. To collect examples of using GSBPM in geospatial statistics production processes, Statistics Finland provides a ten point instruction on how we made our own tests.

In our presentation we discuss the lessons we learnt using GSBPM for describing our geospatial statistics production processes. We also argue that GSBPM with its phases should stay as it is but it's interpretation should be extended by adding guidelines for geospatial statistics dimensions. If we start, in this situation, to point out and stress the specialty of geospatial data (by adding new phases), we may lose a unique chance to consider spatiality as one dimension of statistical data and its important role in the general statistical production process.

Working towards a Geospatial Strategy in Statistics Denmark

Karen Skjelbo (Statistics Denmark, Denmark)

In Statistics Denmark we believe, that we have an unexploited potential for making use of the geospatial dimension in the production and communication of statistics. Therefore, our Board of Directors have established a task force, with the purpose of developing a strategy for making more and better use of Geodata and GIS in Statistics Denmark.

In Denmark there are good digital geographic data, administrative registers with full coverage and good connections and integration between the different registers and the law of Statistics Denmark provides the legal base for Statistics Denmark to get hold of all this data. The basis for increased use of Geodata and GIS in Statistic Denmark is in place.

The new task force consists of staff in Statistics Denmark working with GIS for analysis, GIS for visualization and staff having knowledge about geographical data. To ensure a broad approach in the work with the strategy, the leader of the task force is from the division of "Methodology and Analysis". The work with the strategy is divided in to three tasks.

1. A phased strategy for an increased use of Geodata and GIS in Statistics Denmark. The strategy must contain principles for use e.g. standards, software, data modelling.

2. Specific proposals for new and better use of geography in statistics and analyses.

3. Benefits and costs for the realisation of the strategy.

A part of the work with the strategy is to get input and feedback from users in Statistics Denmark. But it is also considered to invite external experts and users of statistical data and private GIS companies to get inspiration for use and visualization.

The task force is just about to commence its' work. In November the task force will be almost finished with task 1, and I will in my presentation describe the work process and introduce the first draft of the strategy.

Combining statistical and geospatial data – challenges and possible solutions from German official statistics' perspective

Dr. Susanne Schnorr-Baecker, Malte Etienne (Federal Statistical Office of Germany, Germany)

Statistical data consist of a temporal and spatial dimension. This means that they are valid for a certain point in time or period of time as well as for a certain location. At present, official statistics already provides a wide-range of statistical data about social, economic and environmental issues which allows a deep insight in current developments as well as changes over time. Increasingly place based analyses for various units – administrative and non-administrative ones – at subnational matter in politics, economy and science. As a consequence, official statistics is asked to provide interactive maps with statistical indicators which are also included into comprehensive geospatial data-infrastructures (GIS) at national and international level for instance INSPIRE at European level. Additionally, new data sources, techniques and methods in official statistics allow new statistics and statistical information and new forms of collaboration.

Among cartographical mapping for administrative and non-administrative districts, grid-based mapping is becoming more and more important. However, the Federal Statistical Office of Germany had to learn from both the agricultural census 2010 and the population census 2011, which are also available on the internet how challenging grid based data presentation is.

Small area data based on grid-cells increases the range of information of Official Statistics and provides flexible and additional analysis options for users. However, the integration of geography in all phases of the statistical production process also goes along with some challenges to be dealt with of which confidentiality is clearly of major importance. Besides preserving the fundamental legal restrictions which are in Germany predominantly defined by the Federal Statistical Law, it is necessary for dissemination to address the usability of the outcome as well as the data quality. In this regard, considering the criteria and standards defined by European Statistics - the Code of Practice - helps to deal with challenges as it points out which core aspects must be complied with within the production process. This includes – just to mention a few – reliability, accuracy, timeliness and timely, the usage of appropriate methods or the obligation to operate in a cost effective way.

Some major requirements when linking statistical and geospatial data will be described in more depth during the presentation. Furthermore, possible solutions how the respective challenges can be mastered from the German official statistics perspective will be illustrated.

Wednesday,

16 November 2016

SESSION 4: MANAGEMENT OF CENSUS GEOGRAPHY

Improving quality and added value of geo-statistical units

Ignacio Duque (National Statistical Institute, Spain)

Production and maintenance of geographical features of statistical units is a core task for National Statistical Institutes. The accurate boundaries of statistical units are also a mandatory task, because nobody outside NSI could make this work. But, despite its apparent simplicity, it is a tricky task, especially due to the close link with the troubled field of administrative boundaries, from Municipalities to Member States.

The presentation explains the aims and detailed sub-projects developed by the National Statistical Institute of Spain in this field, considering the institutional context, actors and the relationships with other agencies, mainly the Map and Cadastral agencies. The explanation of the goals achieved and remaining during the past five years could be a source of ideas that should be helpful for the statistical production in other contexts.

The first goal of Statistical Units updating is the maintenance of enumeration districts through time and the visualization of them by many tools and services. Nevertheless in the case of the Spanish Statistical System the boundaries of the other statistical set of units (the geographical "Nomenclátor") is also considered. The "Nomenclátor" is little used, but still interesting for its century and a half record of infra-municipal disaggregation.

For enumeration districts updating the main tool is the use of edition utilities inside the viewer of Spatial Data Infrastructure. The work has been done by the appropriate experts of the 52 provincial offices of INE in recent years and the main results are presented: the tools implanted, the workload registered and the experience for improving these tasks.

The key point of the role of municipalities and other agencies as suppliers of different information in the process is explained, and many open questions arise around this subject.

The diffusion of geodata of statistical units inside and outside the statistical office is the final and very important stage of the project. WMS, WFS and others services are considered.

The metainformation about the intertemporal change of statistical units is a key added value of the process and a valuable information to the users outside the statistical office. The INSPIRE recommendations in this issue are not very detailed and a methodological proposal for the production of this information is explained.

Statistics Canada, Temporal and Spatial Opportunities

Elaine Castonguay (Statistics Canada, Canada)

Over the last four censuses of population and agriculture, Statistics Canada developed an integrated approached to manage the development, the maintenance and the dissemination of the census geographic areas. This approach has proven to be flexible without compromising the NSOs ability to conduct longitudinal analysis.

The presentation will provide an overview of the spatial data infrastructure developed to support the maintenance and the dissemination of the census geographic areas as well as key discussion points related to the challenges through time.

Egyptian Experience in integrating statistical and geospatial information

Mennat Allah Anwar Ahmed Hashim (Central Agency for Public Mobilization and Statistics (CAPMAS), Egypt)

The Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS) has made a lot of applications that express how statistical and geospatial information can be integrated and how it can be very useful in decision making.

One of these applications which has been recently published is "Egyptian rural Villages Survey" which shows state of the villages society in all its aspects and also the weaknesses and shortcomings of the existing services in this community as the priorities of the services required of this community, this application also produces different statistical and geospatial analysis that can support decision making in improving the standard of living of these villages.

Another application is the "Century Census" which takes us in a tour to see the history of the Egyptian censuses, starting from census 1882 till the last census in 2006.

Our preparation for the next census 2016/2017 which is based on using digital detailed maps with scale 1:500 covers all Urban and Rural regions, this digital maps will be uploaded on Tablets devices which will be used for the first time in Egypt for collecting the census data.

The next Egyptian census is considered to be a golden chance to spread and publish a new unique numbering and geocoding system for each unit in a dataset, such as a building, household or business based on National Grid system with coordinate system MTMWGS84 (Modify Transfer Mercator).

The next census will help us in accomplishing and fulfilling the five principles of Global Statistical Geospatial Framework which are:

– Use of authoritative geospatial infrastructure and geocoding.

– Geocoded unit record data in a data management environment.

– Common geographic boundaries for dissemination of statistics.

– Interoperable data and metadata standards.

– Common Accessible and usable geostatistical information.

CAPMAS also is going to disseminate all these applications and development steps on its Geospatial portal in the next few months.

SESSION 5: ENVIRONMENT STATISTICS

Compilation of land use statistics using GIS

Erik Engelien, Emil Elisa Traaholdt Vågnes (Statistics Norway, Norway)

In order to produce cost-effective land use statistics covering Norway, existing cartographic databases and registers has been used. Statistics Norway has created a hierarchical classification system, “Standard for classification of land use and land cover”, which is based on both national and international standards and nomenclatures. The method applied is based on utilising the highest quality data sources available, but where no optimal data source exists, the next best quality data sources are used. In practical terms, the method is an automatic geographic information system (GIS) that defines, classifies and assembles the data into a hierarchy. In the statistics on land use and land resources all areas with buildings are classified as built-up. In addition, the building types within a built-up area determine the classification of that area.

Integrated system of natural capital and ecosystem services accounting in the EU – an introduction to the project from a geospatial statistics perspective

Ekkehard Petri (Eurostat, Luxembourg)

The 7th Environment Action Programme (EAP) and the EU Biodiversity Strategy include objectives to develop natural capital accounting (NCA) in the EU, with a focus on ecosystems and their services1. Ecosystem services are the benefits that nature provides to society. Ecosystems services include food provision, air and water filtration, pollination, climate regulation and protection against natural disasters such as flooding and many others. The goal of ecosystem accounting is to include nature and biodiversity and the services they provide in decision making, and to promote more resource efficient and sustainable choices about our future.

The Knowledge Innovation Project on the Integrated system for Natural Capital and ecosystem services Accounting (KIP-INCA) aims to design and implement an integrated accounting system for ecosystems and their services in the EU by connecting relevant existing projects and data. Several Commission Directorate Generals and the EEA work together in this project2. The goal is to build a system that is interoperable with relevant national efforts3.

In a first step KIP-INCA would be designed to produce accounts of: ecosystem extent (delineating ecosystems and measuring changes in areas covered by ecosystems), ecosystem condition (capturing parameters that are linked to essential ecosystem processes, per type of ecosystem); ecosystem service supply and use (connecting the generation of ecosystem services by ecosystems to the use by economic sectors and activities, following international classifications).

Creating the geospatial data platform proposed for KIP-INCA involves taking data sets from a wide range of sources: (1) earth observation (e.g. on land cover), (2) statistical collections including physical data about human activities (e.g. land use, industrial use), biomass production, water use and availability, (3) environmental monitoring data including data reported under relevant legislation (e.g. the Birds and Habitats Directives or the Water Framework Directive) and (4) models which quantify ecosystem services such as water, air and soil regulation, pollination, carbon release and sequestration, etc. and integrate them into a common framework.

A large contribution to the KIP-INCA data platform should come from official statistics with the most obvious one being land use and land cover statistics such as LUCAS and agriculture statistics. However official statistics are often not available at the required spatial resolution and disaggregation of data is required.

The presentation will give an introduction to the KIP-INCA project and present several of the data challenges from a geospatial-statistical perspective.

KEYNOTE

Why does using individual and geocoded data matter in urban economics?

Pierre-Philippe Combes (Lyon University and Sciences Po Paris)

SESSION 6: BIG DATA AND EARTH OBSERVATION DATA IN OFFICIAL STATISTICS

EO4Urban: Sentinel-1A SAR and Sentinel-2A MSI Data for Global Urban Services

Yifang Ban (KTH Royal Institute of Technology, Sweden)

With more than half of the world population now living in cities, and 2.5 billion more people expected to move into cities by 2050, urban areas pose significant challenges on the environment. Although only a small percentage of global land cover, urban areas significantly alter climate, biogeochemistry, and hydrology at local, regional, and global scales. Thus, accurate and timely information on urban land cover and their changing patterns is of critical importance to support sustainable urban development.

EO4Urban is a project within the ESA DUE INNOVATOR III program. The overall objective of this research is to evaluate multitemporal Sentinel-1A SAR and Sentinel- 2A MSI data for global urban services using innovative methods and algorithms.

Based on the user needs, the following specific objectives are put forward: to map urban extent at a global scale, to produce detailed urban land cover maps, to detect new builtup areas, and to map urban green structure as well as its changes at a regional scale. KTH Urban Extractor, a robust algorithm was adapted for urban extent extraction and KTH-SEG, a novel object-based classification method was used for detailed urban land cover mapping. Ten cities around the world in different geographical and environmental conditions are selected as study areas. In the first phase of the project, multitemporal Sentinel-1A IW SAR and Sentinel-2A MSI data over Beijing, Stockholm and Lagos were acquired during 2015 vegetation season.

The results showed that urban areas and small towns could be well extracted using multitemporal Sentinel-1A SAR data. With 10 m resolution, Sentinel-2A MSI data was able to better map low-density builtup areas than Sentinel-1A SAR data. The fusion of SAR and optical data could reduce commission errors of urban extractions from each sensor alone. For urban land cover mapping, the results show that multitemporal Sentinel-1A SAR data alone yielded an overall classification accuracy of 60% for Stockholm. Sentinel-2A MSI data as well as the fusion of Sentinel-1A SAR and Sentinel-2A MSI data, however, produced much higher classification accuracies, both reached 80%. Compared to the urban extraction results from ENVISAT ASAR or ERS SAR data in 1995 and 2005, the urbanization patterns and trends in Beijing, Stockholm and Lagos are being analyzed and will be presented. Urban green structure and their changes are being derived from the urban land cover maps and will also be presented.

Cop4Stat_2015plus – Analysis of COPERNICUS remote sensing data for areal statistical purposes

Stephan Arnold, Thomas Wiatr (Federal Statistical Office (DESTATIS), Germany)

”COP4STAT_2015plus“ is a cooperation project between the Federal Statistical Office (Statistisches Bundesamt, DESTATIS) and the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, BKG) in Germany. The project aims at evaluating the possibilities of the Copernicus remote sensing products for statistical purposes on national level regarding information on land cover (LC) and land use (LU). The existing national and European land cover and land use classification systems use different definitions of classes in comparison to those used by the Statistical Office. A key criterion is the clear and consequent separation between LC and LU definitions. This requirement is manifested in the LUCAS classification system as used by EUROSTAT. The aim of the project Cop4Stat is to identify reproducible and consistent methods to derive LC and LU information from remote sensing data which can be analysed for the

purpose of areal statistical calculations. As input data, multispectral satellite imagery from optical Sensor Sentinel-2a will be used. As additional data sources other existing datasets (national land cover model LBM-DE, Copernicus product High Resolution Layer) are integrated in the work process. The multispectral datasets will first be analysed using established remote sensing methods of segmentation (e.g. watershed segmentation), classification (e.g. supervised classification) and image processing algorithms (e.g. spectral indices). This approach will be processed for selected test areas to finally compare the calculated statistical results from remote sensing data with the official land use statistics (based on cadastral data).

Collaboration to produce official statistics from satellite data products from the Australian Geoscience Data Cube

Martin Brady, Matt Jakab, Richard Dunsome (Australian Bureau of Statistics, Australia)

The detailed classification of land cover from medium resolution multispectral imagery can be challenging when performed at the continental scale, as is required in Australia. A practical alternative is the calculation of Fractional Cover, which indicates the fraction of cover in three broad classes - green vegetation, dry vegetation and bare soil. In Australia, the value of producing official statistics derived from Earth observation data products such as Fractional Cover is being realised. The Australian Bureau of Statistics (ABS) is currently collaborating with Geoscience Australia (GA) to access the Landsat imagery archive for Australia. This archive is stored in the Australian Geoscience Data Cube (AGDC) on the National Computational Infrastructure (NCI) and covers a period from 1987 to the present with a 16 day repeat collection interval. GA is supporting ABS to calculate statistical summaries of Fractional Cover within Statistical Areas as defined in the Australian Statistical Geography Standard (ASGS). Changes in Fractional Cover through time in these areas will be analysed to investigate relationships between land cover, land use and land value across Australia. The results of this work will be featured in upcoming ABS Experimental Land Account publications, and is the first step in a broader collaboration between ABS and GA to leverage time-series Earth observations from a selection of sensors for the production of official statistics for Australia.

Integrating Geography and Statistics, but what about Earth Observation?

Ola Nordbeck (Norwegian Space Centre, Norway)

The on-going digital revolution is providing us with more data and allowing increased and unprecedented possibilities to combine data. This revolution is important in order to follow the global changes in time and space (geographical location) which is crucial for the success of the 2030 Agenda for Sustainable Development.

The EFGS (European Forum for Geography and Statistics) is aiming at making the statistical community to work closer with the Geographical Information (GI) community. This work has progressed through various initiatives allowing these two communities to identify obstacles for better integration.

A third actor, the Earth Observation community, is now investing heavily in new satellite programmes. These investments will result in the provision of frequently updated satellite data allowing the GI and statistical community to get a better understanding of changes over time and potentially resulting in more updated registers.

A collaboration between these three communities is important and has a lot of potentials. Differences in perception and policies are however challenging. It can therefore be tempting to choose alternative avenues for data collaboration, which can have undesirable consequences.

This presentation provides examples of the potentials in collaboration between the three above mentioned communities and recommendations for the way forward.

Mining Mobile Phone Data to Recognize Urban Areas

Stéphanie Combes, Marie-Pierre de Bellefon (Statistics, France), Maarten Vanhoof (Orange Labs – University of Newcastle, United Kindom) and Thomas Plötz (University of Newcastle, United Kindom)

Understanding territory organization, for example in terms of employment, home location and mobility, is crucial for the implementation of policy measures. In France, the National Statistics Office (INSEE) produces a zoning (ZAUER: Urban Area and Rural Employment Area's Zoning) to identify the geographical extent of cities' influence over their environment at the national level. Producing this typology is a complex task. It involves multiple actors and methods, and many arbitrary thresholds have to be chosen. As a consequence a zoning is characterized by long delays between consecutive updates. Recently, mobile phone data has shown promising results for land use classification as they provide for disaggregated, geo-localized and timely information on activity patterns of large shares of populations. In this paper, we exploit a dataset of hourly mobile phone activity profiles collected at each antenna by the French operator Orange to investigate the capabilities of mobile phone data to reproduce the French Urban Area Zoning. Since the ZAUER classification uses commuting information to delineate urban areas, we hypothesize our dataset to be particularly suited for reproducing this zoning by means of supervised classification techniques. In particular, we compare the spatially smoothed predictions of penalized logistic regressions, boosting trees and random forests algorithms using the Fuzzy-Kappa remote sensing metric to account for the fuzziness of our context (in terms of location and cate-gories). Our best results depict an excellent prediction of urban clusters but more difficult disentanglement of rural areas. Besides showing the relevance of mobile phone data for land use classification tasks at a nation-wide scale, our paper explicitely elaborates on the experience of using supervised classification procedure to produce and control the quality of an official statistics indicator.

Global Population Distribution: a continuum of modeling methods

Greg Yetman, Kytt MacManus (CIESIN, Columbia University, USA)

Global population modeling efforts have evolved from simple approaches using census data to geostatistical models that integrate multiple correlates, through to the use of artificial intelligence to disaggregate census population to settlement locations. Additional approaches eschew census data and rely on mobile device and social media streams to model population at extremely fine resolutions in time and space. Inter-comparison of these models and their outputs is lacking. A review of the methodologies used is presented with a discussion of the implications for use (and misuse) of model outputs in analysis and applications.

Merging big data and official statistics for modelling statistical commuting

Pasi Piela (Statistics Finland, Finland)

Commuting distance and commuting time can be calculated as a point-to-point estimate from almost every employed's home to corresponding workplace (the national coverage being about 93 percent). In order to make modelling more realistic different data sources have been merged with current administrative census data at the micro level.

The study to be presented promotes sustainable development by taking into account bicycling (or walking) and public transport opportunities, and specifically the modern urban-rural classification by the Finnish Environment Institute (syke.fi) at micro level.

The speed estimation for a private car usage follows a rather complex estimation structure. Speed estimations for each road element are made by using several variables on a national route network, Digiroad (digiroad.fi). This estimation is enriched by the traffic sensor data by the Finnish Transport Agency (FTA).

Cycle commuting is modelled more straightforwardly in a simplified manner. Commuting model is based on a shortest non-hierarchical path including separate cycle paths and excluding motorways (according to the traffic regulations). The average speed is assumed to be 17 kilometers per hour.

The estimation of the public transport accessibility is implemented by utilizing the open application programming interface (API) of the Journey Planner of the Helsinki Regional Transport Authority. Correspondingly FTA's country-wide public transport Journey.fi service has been researched and the results will be presented.

Results show clearly differences by region and area type. The presentation also discusses the data safety issues on a modern computing environments.

SESSION 7: GEOSPATIAL ANALYSIS BEST PRACTICES 1

Exploring Europe's border areas with the help of geostatistics

Hugo Poelman (European Commission - DG Regional and Urban Policy, Belgium)

The EU has a longstanding tradition of cross-border cooperation programmes, operating in all border areas of the EU.

Cross-border cooperation programme areas are defined as groups of NUTS3 regions. While this approach has obvious regulatory and implementation advantages, the assessment of needs, challenges and opportunities of border regions encounters some obstacles of data availability at NUTS3 level. In addition, the substantial size differences amongst NUTS3 regions make comparisons between border areas somewhat problematic.

The use of grid-, point- and network-based data helps to alleviate these obstacles. Firstly, grid data have been used to determine the population distribution along borders, in areas independent from the NUTS boundaries. This process has not only covered terrestrial borders, but has also been applied to refine the analytical definition of maritime border areas.

This analysis leads to a classification of NUTS3 regions according to their share of border area population.

Subsequently, the border areas can be combined with other geodata. For instance, border areas are being characterised in relationship with a classification of mountain areas, road accessibility between selected settlements can be assessed, as well as rail accessibility between all stations located in border areas.

These analytical outcomes, together with NUTS-related analysis, are expected to feed the preparation of the future EU territorial cooperation policy.

Fly across the boundary: A tracing of actual movements in air transportation network

Dongkun YIM, Suyoung KANG (Department of Geography, Seoul National University, South Korea)

Interpreting flow statistics are intellectually challenging tasks in various fields. As regards to mobility and migration, geography also deals with the different phenomena of flows. In this paper we discuss the statistics on the movements across the boundary. As a kind of "boundary problem", statistics on intra-regional mobility or migration status reveal only a small part of the actual movements. How does it reflect the actual movements, the part of flows passing through the boundary? Especially the frontier with a few gateways draws the flux into several paths, so that the spatial interactions between the regions are simplified to the relation of great poles. We will review this issue through the prism of the air traffic from and to Asia. In an air transportation network, a few airports support the intercontinental flights and the movements between the cities in different continents might be veiled. However, considering that air traffic data have the information of origin and destination of the movements, we can investigate the relation of the segment of flows over the boundary to their actual movements between the regions. In turn, we ask the conditions in that flow data over boundary represent the actual movements.

Income levels and inequality in metropolitan areas: a comparative approach in OECD countries

Justine Boulant, Monica Brezzi and Paolo Veneri (Regional Development Policy Division, OECD, France)

This paper assesses levels and distribution of household disposable income in OECD metropolitan areas. All indicators were produced through a dedicated data collection, which, for most countries, uses administrative data from tax records available at detailed local scale (i.e. municipalities, local authorities, counties, etc.).

Using different estimation techniques, we provide internationally comparable figures for 216 OECD metropolitan areas. The results highlight stark differences in both income levels and inequality within metropolitan areas, even for those belonging to the same country.

These new data allow to benchmark how metropolitan areas in OECD countries fare on many SDGs and the size of within country inequalities that may hamper reaching the SDGs.

Geospatial Data and Statistics at Austrian Post Office

Karel Mauric, Simone Ortner (HQ of Austrian Post, Vienna, Austria)

Austrian Post is the leading logistics and postal services provider in Austria. Its main business activities include the transport and delivery of letters, direct mail items, print media and parcels. The branch network of Austrian Post ranks among the largest private customer networks in the country, offering high-quality postal, banking and telecommunications products and services to its customers throughout Austria. The company makes an important contribution to safeguarding the nation's communications and logistics infrastructure based on its nationwide and reliable supply of high quality postal services on behalf of the Austrian population and economy

Geospatial information and statistics are strongly integrated in many important business solutions and internal services at the Austrian Post. Firstly, our presentation will give a short overview of the main geospatial applications with representative examples of involving/implementing statistics and demographics in the daily work of many departments of the Austrian Post. Then, using the example of a Geomarketing-Solution, the usage of statistics data will be demonstrated in more detail.

Geomarketing lets advertisers address their desired target audiences with pinpoint accuracy and without any wastage. You can define your target audiences by selecting age, purchasing power, city/town of residence, size of city/town and many other criteria. Austrian Post will then select the delivery areas and locations that fit your distribution plan and your previously defined target audience. Our database makes it happen, for it matches geographical data (such as delivery areas and postcode areas) with sociodemographic criteria (age, marital status, children, etc.) as well as socioeconomic criteria (including purchasing power).

Benefits of using Geomarketing Services

- Increased response rates
- Reduced wastage
- Saves money thanks to fewer printed materials and less postage Geomarketing is a great choice, whether you would like to address a clearly defined target audience or all people in a defined area surrounding your business. Following options are available.
- Target audience distribution plan This option lets you define distribution areas by selecting from approximately 40 criteria including age, gender, number of children, level of education, purchasing power, etc. per area/household/person.
- Postal branch distribution plan

Relying on geomarketing to create a distribution plan based on our list of postal branches. You can select any area surrounding the postal branch of your choice, thereby determining the number of recipients. Upon request, you also have the option of adding target audience specific criteria.

Thursday,

17 November 2016

SESSION 8: GEOSPATIAL ANALYSIS BEST PRACTICES 2

Population distribution in the Republic of Kosovo: a comparative analysis on urban population and its classification based on administrative and non-administrative criteria

Idriz Shala, Burim Limolli (Kosovo Agency of Statistics, Republic of Kosovo)

Statistical data for urban and rural areas are of some considerable importance for the central government and for local authorities while planning and managing services for local communities. For instance, the allocation of health and social care funding, housing, roads, water and sewerage and the provision and maintenance of schools have all distinctive aspects in urban and rural areas. Employment for urban and rural population has different features as well.

Recently, in many countries, including the Balkan countries, this distinction has become unclear and the principal difference between urban and rural areas in terms of the circumstances of living tends to be a matter of the degree of concentration of population. Indeed, rapid urbanisation processes have greatly raised the need for actual information related to different sizes of urban areas, and to the need to define standards for data comparability.

Like in all countries conducting a population census, in Kosovo the census data was disseminated following the administrative structure of the countries based on census legislation. In most of the Western Balkan countries, the law classifies the administrative units as urban or rural. However, like in many other countries, their cities and other urban areas are usually enlarging their size faster than the capacity of the law to revise such definitions which are needed to make urban boundaries consistent with the actual size of urban and non-urban areas. Therefore, the breakdown of census data by urban areas, both at national and regional level, is underreported. As a consequence, census results show significant differences in terms of urban/rural breakdown if different criteria for data classification are applied. Moreover, taking into account that the 2011 censuses have been used to update the sample frame for household surveys, also their results are affected by the definitions used for urban and rural population.

The main objective of this paper is to analyse the distribution of the urban resident population of Kosovo, as obtained from their 2011 population and housing censuses, according to administrative criteria and on the basis of a new approach for data classification as well, and to compare the differences at national and regional levels. Selected census variables are tabulated by urban/rural modalities using administrative and non-administrative criteria, including the 1 km² grid-based typology recently adopted by the European Union.

Official grid-based statistics: regional statistics in Andalusia (Spain)

Iria Enrique Regueira (Institute of Statistics and Cartography of Andalusia, Spain)

The Institute of Statistics and Cartography of Andalusia (IECA), official body member of the Regional Government of Andalusia, is the result of the merge in 2011 of the Statistical Office of Andalusia and the Institute of Cartography of Andalusia. The IECA has worked ever since on the integration and synergies of these two areas: Statistics & Geography.

Namely, IECA is engaged in a long-term project to integrate the grid of cells sized 250_250 m as a standard in statistics production procedures and data dissemination. Hereunder we share some of the milestones, lessons learnt and future goals of this project.

In 2013 “A population grid for Andalusia. Year 2013” was published, a first hybrid approach to population georeferentiation (top-down & bottom-up). A second edition of “A population grid for Andalusia Year 2013” was released in 2014 improving georeferentiation, thanks to 2011 building census data, and achieving 97.2% bottom-up georreferentiation of the population settled in Andalusia. This edition built entirely from a bottom-up approach provided us with a benchmark to review and validate the initial hybrid approach. Lessons learnt from 2013 grid editions and improvements in GIS infrastructure integration at IECA lead our current work on population grids. A shared PostGis Infrastructure is being built in order to share and consult updated unique data (population register, cadastre, census…) and initial linkage results for 2014 population register attained georeferentiation of 94% of populated buildings, standing for 95% of the population settled in Andalusia on 1st of January 2014.

Furthermore, since 2013 first edition, population grid data has been required and proved to be highly valuable for administrative planning & evaluation, namely regional educational planning or infrastructure planning. Additionally, grid standard has also been used to elaborate derived indicators such as Bicycle lane accessibility or Smoothed Standardized Mortality ratio by grid Cells, recently published by IECA.

Finally, IECA has also initiated georeferentiation of establishments, in order to provide a supplementary view of disaggregated activity and employment in Andalusia.

Capturing the Synergy of Geospatial and Statistics : A Singapore’s Perspective

Ng Siau Yong, Angelinie Winarto (Singapore Land Authority, Singapore)

Globally, it is acknowledged that the integration of statistical and geospatial information aid to improve the applicability of evidence-based decision-making.

Statistics are used to understand and make decisions on complex economic, social, security and environmental issues. Spatial information is crucial to derive patterns not readily apparent to the observer by using Geospatial Information System and Technology (GIST). However, to get a holistic view of the economy, society, security and environment, integrating spatial information with statistical information and processing the data by spatial analysis methodology is essential.

Through the spatial-analysis approach, and methods of coordination, harmonisation and collaboration between statistical and geospatial establishments, the observer is able to make sound decisions by finding clarity from complexity, to tackle some of the world’s most pressing challenges (i.e. transportation, security, healthcare and housing).

In 2012, the Secretary General of the UN Economic and Social Council stated: “The work on global geospatial information management over the past two to three years has confirmed that one of the key challenges is better integration of geospatial and statistical information as a basis for sound and evidence-based decision-making”.

Some Department of Statistics in the world use GIST for the purpose of census and demographic mapping and for updating census and demographic maps. There is, however, an increased acknowledgement that spatial analysis is important not just for national statistics, it provides a structure for collecting, collating, processing, storing, aggregating, generalising, disseminating and analysing operational and business data.

Tobler’s First Law of Geography states that “everything is related to everything else, but near things are more related than distant things”. We should be mindful of the power of location and statisticians are

aware that the harmonisation of statistics, location and boundary delineation could affect sampling and therefore their analysis and results. By linking people, business, economy and environment to a location, the outcome is a better understanding of the social and economic issues of that location.

Today, there is an increased demand for geospatial analysis of the socio-economic data by various users. By integrating geospatial and statistical data and processing the data via spatial analysis, it facilitated the building of complex multidimensional location-based information resources. This could potentially generate unique visualisation, insightful spatial analysis and valuable predictive modelling results in the form of a map. The integration of statistical and geospatial data lead to cost savings, greater credibility, increased accuracy and better decision-making by the various users.

This presentation aims to share Singapore's perspective on the postulations above. Experience in working with various public agencies in integrating geospatial and statistics for service delivery and policy decisions will be presented. Some Asian examples of the use of geospatial and statistics will be discussed.

Enhancing reliability of soil sealing indicators by use of geostatistical modeling

Patrick Sillard (Observation and Statistics Service - Ministry of environment, France)

Soil sealing (imperviousness) indicators are among most important indicators used to follow the extension of soil consumption. They are useful to monitor the pressure of human activity put on soil and at the end on biodiversity. Various sources are available to compute Soil sealing indicators: CORINE Land Cover and especially the High resolution layers, LUCAS survey, highly accurate geographical databases such as the property tax geographical database. All these data are acquired with different methods of observation and at different working scales. They lead, without further analysis, to very different rates of imperviousness. For example the rate of imperviousness estimated from the CORINE land cover geographical database is 50% less than the one we may compute from LUCAS.

The goal of the paper is to develop a statistical model of the imperviousness seen as a continuous spatial stochastic process that makes it possible to conciliate the various observations. As an application, the High resolution layers of CORINE and the LUCAS Survey are compared on the French territory. In particular, we study the consequence of the specific structure of the process (autocorrelation, long memory) on the imperviousness indicators and their associated variance.

A Room with a View or Rear Window? The Demand for Housing Attributes in Paris

Mathilde Poulhes (French Ministry of the Environment, Energy and Marine Affairs)

This paper estimates buyers' preferences for dwelling attributes and neighbourhood characteristics. The collected data allows for the simultaneous consideration of a wide range of intrinsic characteristics, such as building type, surface, floor, etc., and environmental characteristics, including noise, crime, school quality, distance to jobs, etc. The marginal willingness to pay is identified from transaction data under the hypotheses of the hedonic model described by Rosen (1974). Estimation is achieved by using flexible semi-parametric methods. Characteristics explain more than 90% of the variance of dwelling prices, showing a positive marginal willingness to pay for job accessibility and school quality and a weaker but significant negative marginal willingness to pay for a higher crime rate in the area. By contrast, noise level or public transport accessibility have less influence on housing prices. These results are robust to the inclusion of census tract fixed-effects, which also drastically reduces the spatial correlation of the residual prices from 0.10 (significant at 1%) to an insignificant correlation of 0.001.

SESSION 9: DISSEMINATION ISSUES

Geostatistics Portal – new developments and future plans

Miroslaw Migacz (Central Statistical Office, Poland)

Developed since 2011 and officially launched in 2013, the Geostatistics Portal is the place, where statistical data users can find everything they need to visualize data on thematic maps. Though the system was initially built for disseminating census results, the data scope and tool variety is growing constantly since the launch.

The Local Data Bank – a database with vast amounts of statistical data updated on monthly basis is now fully accessible via the Portal. Thematic presentations can be prepared for data since 2002 and presented phenomena can be compared within a selected set of years. If this extensive database is not enough, users can import their own data and visualize them using the complete set of thematic map tools.

The Geostatistics Portal offers a broad choice of geovisualization tools. Customizable choropleth maps can be created to visualize statistical indicators, whereas absolute data can be shown using a powerful set of diagram map tools. Single and multiple phenomena can be presented using different kinds of diagram maps. Users can visualize absolute values, reflect a structure of a phenomenon or show its trend across a span of several years. Complex diagram maps allow presenting data in different units and different scales on one single map. Any thematic map created in the Portal can be printed out or saved as a document, while statistical data can be exported as a table.

While most data banks and portals offer data for basic administrative and statistical units (like NUTS and LAU), the Geostatistics Portal goes a few steps further. Statistical grids for the most popular demographic classifications are available for view and download. If those static presentations are not enough, users can perform queries on census microdata drawing a freehand polygon. An aggregated result of such a query is returned to the user, provided the statistical confidentiality rules are met (a check is performed “on-the-fly”).

At the moment, the Geostatistics Portal offers access to an extensive set of statistical data, advanced visualization tools, data presentations on a level lower than LAU2 and mechanisms for microdata querying on any desired area. A set of tools and presentations is also available in a mobile version of the system (available as an application). Furthermore, the Portal is also the Polish statistics’ gateway to INSPIRE datasets and services, which are fully compliant with the Directive’s technical guidelines.

The already impressive set of tools is to be developed further. Future plans include introducing georeferenced linked open data, exploratory spatial analyses, geostatistical modeling and semi-automated enrichment of user supplied data. The existing tools are to be updated, expanded and made more flexible – all this to bring statistical data closer to the people and authorities and make governance easier on all administrative levels.

An open and closed case? Dissemination and the use of proprietary suppliers in an open data world

Ian Coady (Office for National Statistics, United Kingdom)

Statistical organisations are increasingly challenged to make use of open data tools and software to support the integration of statistics and geography as part of a wider shift by public sector organisations to move away from the high cost and bespoke development of commercial suppliers. ‘Open’ has its own challenges however with the need to build enough capability and expertise to be able to support a corporate approach to open software.

This presentation will look at a case study from the UK where the NSI has worked with a commercial supplier (ESRI) on the delivery of a portal that can meet the demands of open data.

It will examine the challenges of publishing high quality geographic data and how the shifting demands of the user community is changing the methods through which we make our data available. Finally, it will look at how the statistical-spatial community could develop generic shared platforms to support a common approach to statistical production such as that set out by the Common Statistical Production Architecture for statistics.

Table Joining Service: The solution for INSPIRE themes without geometry

Peter Bresters (Statistics Netherlands, Netherlands)

A Table Joining Service (TJS) is an online service that links online statistical tables to map services with temporary online thematic map services as a result.

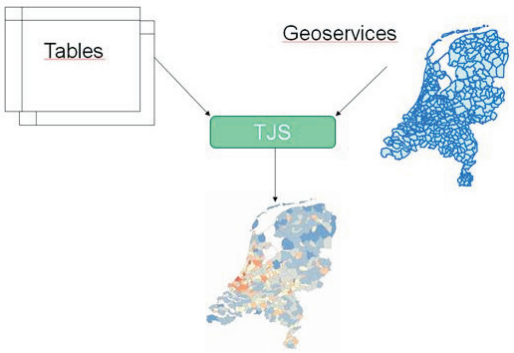
This presentation shows the impact analysis of a Table Joining Service (TJS) in the national infrastructure that the Netherlands uses for the European project INSPIRE (PDOK). It is the final delivery product of the Eurostat Grant awarded to Statistics Netherlands in cooperation with the Dutch Cadaster, PDOK and Geonovum. It is created on the bases of a proof of concept of the TJS as executed by these organizations. This proof of concept has shown how a TJS can work for the machine readable open data from Statistics Netherlands, combined with a map service with open topographic data for statistical units. A demo will be presented.

One of the strongest motivations behind the TJS is the European project INSPIRE, but other motivations are also mentioned in the presentation.

The impact analyses not only looks at the Dutch situation but also gives some thoughts about what impact it might have for the rest of Europe.

One of the conclusions is that a central European TJS could be a very cost effective way to realize the INSPIRE goals.

Finally some lessons learned and recommendation are described.



Spatial Statistics on Web 2 and Oskari

Timo Aarnio, Jani Kylmäaho (National Land Survey of Finland, Finland)

Statistics Finland and National Land Survey of Finland (NLSFI) will again join forces in the collaborative project Spatial Statistics on Web 2 (SSW2). In the previous SSW-project NLSFI developed analysis functionality in the open source Oskari software (www.oskari.org) in order to add value to the spatial and statistical data made available by Statistics Finland as web services. Oskari is a widely used and feature rich web mapping platform with support for statistical mapping and spatial analyses. In the sequel project SSW2 the emphasis on Oskari development will be on the thematic mapping functionality.

The current version of thematic mapping in Oskari was developed to utilize a single existing web service that provides statistical data. In SSW2, the aim is to make the functionality a lot more generic so that multiple sources of statistical information, such as SDMX REST 2.1 and some other REST/JSON based services are supported. We also plan to introduce additional types of visualizations such as charts, graphs, animations and different types of thematic maps.

One of the core functions of Oskari is map publishing, a tool that lets the end-user create a map application without any programming expertise. The resulting map application can then be embedded onto any web site. Most of the tools and functionalities of Oskari can also be used in embedded maps. For example, user can create a thematic map application complete with classification tools and statistical data table and embed it as a part of a website. The website might be e.g. an official report, a blogpost or a piece of news.

In the presentation we will dive into the upcoming developments in more detail and showcase a proof-of-concept implementation combining Eurostat statistical data with multiple statistical units. If time permits we can also demonstrate the current and development versions of the thematic maps functionality in practice.