ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

ECE/CES/BUR/2018/OCT/3
27 September 2017

Meeting of the 2017/2018 Bureau
Geneva (Switzerland), 14-15 October 2018

For discussion and recommendations

Item 2 (b) of the Provisional Agenda

IN-DEPTH REVIEW OF SATELLITE IMAGERY / EARTH OBSERVATION TECHNOLOGY IN OFFICIAL STATISTICS

Prepared by Canada, Mexico, Austria and Eurostat

This in-depth review examines the use of satellite imagery and earth observation technology in official statistics. ***The Bureau is invited to discuss the issues and challenges in this area and consider how to address them.***

Table of Contents

I. Executive summary

1. Given the steady advancement of Earth Observation (EO) technology and infrastructure over the past decades, and the selective use of EO data by some National Statistical Offices (NSOs), the objective of this CES in-depth review is to report on the state of NSO activity with respect to this data source, and its potential to address current or emerging needs of official statistics.

2. As context, the work of three important international organizations and three recent comprehensive studies on the actual and potential applications of EO data to various statistical programs are discussed and implications suggested. This is followed by a review of EO-related activities within the statistical programs of Austria, Canada, Mexico and Eurostat.

3. The country review found that EO inputs are commonly used to support agricultural statistics and environmental accounts, with an increasing level of activity in the area of sustainable development indicators such as land-use, climate change, water stress and water quality. In terms of research on emerging applications, knowledge is advancing on how to store and manipulate EO imagery more efficiently (Australia, Mexico), for use in real-estate valuation and maintenance of statistical registers (Austria, Canada), and for the purposes of reporting on SDG-related indicators (Canada, Mexico, Eurostat.)

4. Several cautionary **lessons can be learned** with respect to the use of EO data by NSOs:

- NSOs often lack the expertise in how to process and apply EO data to their needs;
- There is a risk of over-confidence that EO data is a "complete" solution for NSO needs in a given area;
- Transformation of business processes to accept EO input requires time, budget and institutional commitment.

5. Despite these important considerations, the review found real **opportunities for NSOs** in this area:

- The strength of EO data is on macro-level collection and reporting. Agriculture and environment statistics are well-suited to this level of reporting;
- There is growing attention on the Sustainable Development Goals and the indicators supporting them. EO data are well-aligned to a significant number of the target indicators;
- The complexity and cost involved in leveraging EO data as an input to NSO programs provides a compelling reason to collaborate on the development and implementation of common approaches.

6. To the extent that the international NSO community is being asked to play a role in SDG reporting in the future, the authors of the paper identify the following **recommendations** for discussion:

- NSOs should use the SDG framework to drive collaboration on EO-based research. This will accelerate decision-making on which indicators are best supported by EO data;
- NSOs should collaborate on a generalized approach to using EO-data to support the production of specific SDG indicators to reduce costs of processing and integrating EO-enabled SDG reporting;

- NSOs, as a community of practice, should collaborate with the national / international EO community to articulate and implement a generalized approach to EO data collection and processing as input to SDG reporting. Such a development would go a long way to bringing EO-inputs to a level comparable to many of the other inputs used currently by NSOs.

II. Introduction

7. The Bureau of the Conference of European Statisticians (CES) regularly reviews selected statistical areas in depth. The aim of the reviews is to improve coordination of statistical activities in the UNECE region, identify gaps or duplication of work, and address emerging issues. The review focuses on strategic issues and highlights concerns of statistical offices of both a conceptual and a coordinating nature. The current paper provides the basis for the review by summarising the international statistical activities in the selected area, identifying issues and problems, and making recommendations on possible follow-up actions.

8. Developed first for use in military and mapping applications, the uses of satellite imagery have slowly expanded over time; their key strength being the ability to receive and process generalized data for a large area at relatively low cost. In a sense, satellite data can be seen as one of the first sources of 'big data', large-scale datasets requiring substantial processing and storage capacity in order to exploit.

9. Over time, the variety and accuracy of satellite sensors improved, extending the domains of practical application, especially in agriculture, forestry, meteorology, and geology. As NSOs expanded their programs in these same areas, the use of Earth Observation technology as a data collection instrument also began to be considered.

10. Today, all NSOs face an incremental demand to produce more for less. This means more domains and more disaggregation in less time and lower cost. At the same time, they must exploit large-volume datasets that help describe macro-level phenomena such as land-use changes, water-stress, climate change, etc.

11. In order to promote discussion and efficient exploitation of this data source, the CES Bureau selected the topic of use of satellite image data in official statistics for an in-depth review at its October 2018 meeting. The statistical offices of Mexico, Canada, Austria and Eurostat expressed interest in collaborating on a paper to provide basis for this in-depth review.

12. Sections 3-6 of the report outline recent relevant contributions in this area, followed by a summary of the different uses to which satellite data / earth observation data are put by NSOs contributing to this report.

13. Sections 7-9 of the report consider emerging concerns for NSOs with respect to EO technology as a collection instrument, suggest domains where satellite / earth observation technologies might be profitably applied, and recommend that NSOs actively collaborate to accelerate the development of techniques to exploit such data as input to evidence-based decision-making at many levels.

III. Scope / definition of the review

14. The intent of this review is to survey how various types of satellite data and the techniques used to process or analyze them are used to support the Generic Statistical Business Process Model (GSPBM) as applied within National and International Statistical

Organizations. Situating the review in this way helps clarify when, why, and how the data are used within that process.

15. Second, the report is structured so as to make clear the distinction between uses that support existing programs leading to published official statistics, versus research or experimental work. This is important, as it helps to clarify domains in which existing earth observation technology has been proven, versus emerging opportunities or experimental work.

16. Deeper review of the different applications will, at a subsequent phase, improve the understanding of the "business case" for earth observation technology as a data collection / processing mode. More specific questions about the relative costs, quality and timeliness of such data, as compared to other approaches need to be fully understood, as well as any significant limitations of the data or techniques. Given the investment involved, it is also important to consider whether satellite data are being acquired and managed as part of a systematic input to a NSO-wide geographic/geospatial data framework, or whether they have more limited application to specific programs. The "Enterprise Geospatial" approach helps to achieve economies of scale for the NSO as a whole.

IV. Overview of recent activities in the area

17. The field of Earth Observation technology has seen progressive development for several decades. Over time, technological advancements have allowed spatial resolution and refresh intervals to increase, while prices have declined. This, in turn, has spurred work to identify new applications of the data emanating from this infrastructure. EO is an active field, though somewhat less well utilized within official statistics.

18. There are many agencies involved, both private and public, but the efforts of three groups are prominent in advancing the importance of EO data for official statistics:

Committee on Earth Observation Satellites (CEOS) - Established in 1984, this group encourages international collaboration on space-borne EO missions for the benefit of society. Its membership is comprised of more than 30 of the world's space agencies which together operate more than 100 space missions. CEOS has recognized the fundamental role EO has to play in monitoring SDGs, and has agreed to coordinate CEOS agencies providing satellite data for the 2030 Agenda for Sustainable Development (CEOS, 2018a).

Group on Earth Observations (GEO) - Established in 2005, the Group is a network of public and private interests advocating to make EO data more broadly available as a public good (GEO, 2018). This group is also positioning its involvement in the area of SDGs (EO4SDG, 2018).

UN Committee of Experts on Global Geospatial Information Management (UN-GGIM) - Established in 2011, the Committee sets the directions for the production and use of geospatial information within national and global policy frameworks, and for building and strengthening geospatial information capacity of nations, especially of developing countries. This organization also has an SDG sub-committee (UN-GGIM, 2018). Under the auspices of this organization, Mexico, Canada, and other countries have explained the importance of aligning the functions of official statistics with the geo-spatial and EO expertise found in national mapping agencies, and very recently, Statistics Canada and Natural Resources Canada put in place a Memorandum of Understanding to make such collaboration explicit (Rancourt and Shukle, 2017).

19. In addition to the above, three significant documents have been produced in recent years that are germane to the current review, and are summarized below. Collectively, they

provide an overview of the recent state of the art in EO, and suggest some promising emerging directions for its use by NSOs.

A. Handbook on earth observation for official statistics

20. This handbook was produced by members of the United Nations Task Team on Satellite Imagery and Geospatial Data and was released in December 2017 (United Nations, 2017). Its purpose is to guide NSOs which have limited experience in producing official statistics from EO data. There are sections on: EO sources, methodologies, pilot projects, case studies, and a framework for assessing the benefit of EO data implementation in a statistical production process.

21. One key message for NSOs is that the utility of EO data use in a statistical production process should be assessed on a case-by-case basis. A second is that EO data have been identified as an essential data source for NSOs to measure, monitor, and report on SDGs for the 2030 Agenda for Sustainable Development.

1. Case studies of EO in official statistics

- In Australia, the Department of Science, Information Technology and Innovation (DSITI) has utilized time-series Landsat imagery to classify Crop and Non-Crop areas using predictive classification algorithms trained on field-based and image-interpreted training data. An expanded mapping approach has shown good results in mapping broad crop types (Pringle et al. 2018);
- Some non-NSOs produce official statistics from EO data and have production processes similar to GSBPM. Examples highlighted include: United States Department of Agriculture and their monthly *World Agricultural Supply and Demand Estimates* (WASDE) report, European Commission and the *Monitoring Agricultural Resources* (MARS) bulletin, and others.

2. Key messages

- NSOs are encouraged to engage with the EO expert community such as international organizations, national space/science/mapping agencies, academia and private sector to build their capacity to produce official statistics from EO.
- EO data have the potential to meet many key objectives of statistical agencies such as timelier outputs; reduction of survey frequency, costs, increased spatial resolution, etc.
- Issue – EO data processing requires specialist knowledge. NSOs lack expertise in processing and applying such data as input to official statistics.

B. Earth observation in support of Sustainable Development Goals

22. This report was produced by the Committee on Earth Observation Satellites (CEOS) and was released in early 2018 (CEOS, 2018b). Its purpose is to inform all sectors of society how geospatial information, EO data and other data sources can contribute to the 2030 Agenda for Sustainable Development. The report explains the role of EO in support of the SDGs, provides examples to demonstrate some of the possibilities, and then focuses on specific SDGs where EO data are central to quality reporting.

23. The global indicator framework developed for the Agenda 2030 contains 17 SDGs, with 169 Targets and 232 Indicators. EO data has been identified as an essential data source for around 40 of the targets and 30 of the indicators (some indicators are used to measure several targets).

1. Examples of EO in official statistics

24. Earth observation data has been an important source of information for the System of Environmental-Economic Accounting (SEEA) Central Framework and SEEA Experimental Ecosystem Accounting (SEEA EEA). EO derived environmental information is combined with economic statistics to report on the contributions of the ecosystems and the impacts of economic activities on the ecosystems, and to SDG indicators such as land degradation, freshwater ecosystems, and land use. The report highlights approaches used in Australia, Brazil, Mexico and Nepal.

25. Two interesting extensions of these ideas are noted. Statistics Netherlands derived an evapotranspiration (AET) map from EO data. The measurement of AET is important for the measurement of water-related SDG indicators. Second, Mexico's National Institute of Statistics and Geography (INEGI) produced a Collaborative Site for Disaster Response to be used by several government agencies. The platform uses data from satellite optical and radar sensors, drones, and private providers to produce disaster-related data in emergency response.

2. Research to expand use of EO in the GSBPM

26. The document provides an extensive overview of the public and private sector organizations involved in developing, producing, or using EO data, and on their perspectives on its application to SDGs. The review includes a summary of efforts to develop EO methodologies, technical assistance, collection and dissemination of data, and engagement activities to promoting the use of EO within the official statistics community.

3. Key messages

- The report asserts that NSO success in reporting on progress of the SDGs will require national statistical systems to be modernized by supplementing traditional statistical methods and data sources with new sources of data, including geospatial information and EO, as well as advanced data processing and big data analytical techniques;
- Of 30 indicators where EO data has a role to play, only 12 have established methodologies and regular data production by countries, suggesting that there is still significant unrealized potential for EO data to contribute to the global SDG indicator framework;
- Some NSOs may have more difficulty with the institutional and technical challenges related to using EO data for the SDG reporting. Effective partnerships between the major actors in this space will be an important enabler in these cases.

C. Handbook on remote sensing for agricultural statistics

27. This handbook was produced in 2017 by a group of senior international experts in the field of remote sensing (GSARS, 2017). Its purpose is to provide guidelines on how to

assess the utility of remote sensing for agricultural statistics, and how EO can be integrated into GSBPM.

28. As a general statement, the domain of agricultural statistics has seen the greatest application of EO technology in collection and analysis as part of official statistics. The report addresses EO for the production of agricultural statistics in areas such as land cover, field-level annual crop mapping, crop area estimation, Early Warning Systems, crop yield forecasting, and list and area frame design. Other chapters discuss technical aspects related to data sources, software / hardware, and organizational capacity, and assessment of the business case of EO solutions.

1. Examples of EO in agricultural statistics

- The Global Food Security - Support Analysis Data Project (GFSAD30), led by USGS, uses multi-sensor EO data and other sources to produce consistent and unbiased estimates of global agricultural cropland products;
- The ESA Sentinel-2 for Agriculture Project is an open-source system to process Sentinel-2 and Landsat-8 imagery that delivers multiple products generated frequently during the growing season for agricultural monitoring and statistics purposes, and has been successfully demonstrated at national scale for selected countries and other test sites around the world;
- India uses space, agro-meteorology and land-based observations (FASAL) to forecast crop production for 11 major crop types at a district, state and national-level;
- The Chinese CropWatch monitoring system uses EO and ground-based indicators such as pest, disease and agroclimatic conditions to monitor global production. Crop condition bulletins are produced quarterly, for regions, countries and sub-nationally;
- The FAO Global Forest Resources Assessment (FRA) provides an estimate of forest change at five to ten year intervals based on sampled areas and using moderate-resolution imagery. Monitoring forest cover and deforestation is important for reporting on SDG indicators.

2. Issues / opportunities

- The cost-effectiveness of EO for agricultural statistics has been improved with the decrease in data costs, increasing quality of imagery, and open-source GIS and remote sensing software, as well as access to remote cloud processing tools such as Google Earth Engine;
- EO technology, though a viable approach in most areas, still faces some constraints in areas with small field sizes, persistent cloud cover, mixed and intercropping practices;
- Opportunity – A recent trend is to move large data storage and processing to cloud solutions. This approach may provide some advantages to setting up, maintaining, and investing in in-house storage and computing. However, cloud-computing might raise other issues related to licensing and intellectual property rights.

V. Overview of recent activities in selected countries / agencies

29. This chapter provides an overview of the work of selected countries in the statistical area under review. It is assumed that readers have a general familiarity of the goals and internal workings of NSOs, particularly within the context of the Generic Statistical Business Process Model (GSPBM), and that this serves as the context for the review of how such organizations assess the value and exploitation of EO data. Each section is structured to make a distinction between statistical programs currently using EO inputs, as opposed to research initiatives that seek to further expand uses.

A. Eurostat

1. Input to officially published statistics

30. The use of satellite and earth observation data has not yet found its way into the major statistical programs of Eurostat, as the mandate of this multilateral agency is to aggregate national statistics to harmonised European products. So far, the national statistics compiled into EU official statistics come from traditional data sources such as censuses, surveys or administrative data. EO data at this point only play a marginal role in the production of official EU wide statistics.

31. One notable exception to the above is that Eurostat conducts the Land Use and Land Cover Survey (LUCAS) every three years. EO input is essential to that process. LUCAS ortho-imagery and satellite imagery are used for preparing the sample, supporting surveyors in the field work, and for checking the quality of the information collected for individual sample points.

2. Research and development

32. As part of a wider research agenda on the use of big data for statistics, Eurostat is investigating some areas of sensor data which are at least similar and complementary with the use of satellite data:

<u>Use of Automatic Identification Systems (AIS)</u>

33. In the marine sector, AIS is used to ensure that ships at sea can be identified and monitored during their passage, primarily for reasons of safety and traffic control. Eurostat is experimenting with the use of this technology as an input to transport statistics. Several EU member states have conducted similar projects for their national statistical systems.

<u>Sustainable Development Goals</u>

34. For the set of SDG indicators that Eurostat has specifically designed for the EU, Eurostat, together with other EU services, is investigating the use of Copernicus and other satellite mission information, mainly for the collection of statistics on land degradation (soil-sealing and land take). High resolution mapping is used indirectly to calculate urban indicators such as access to green areas in cities. This is a promising approach to produce the related SDG indicator on access to open space.

<u>Agriculture / Aquaculture</u>

35. Other areas of the European Commission are heavy users of EO data as a supporting source for statistics on crop yield, monitoring fishing boats/quota, or the planning of

aquaculture installations. Several European NSOs have started exploring EO data for agriculture statistics, supported with grants issued by Eurostat.

Greenhouses / Solar panels

36. Eurostat also supports financially research activities in NSOs, e.g. in the Netherlands to detect sealed areas or solar panels.

B. Mexico (INEGI)

1. Input to officially published statistics

37. At a general level, INEGI has produced maps of Mexico for several decades; initially in paper format (1970s, 1980s), and now in the form of themed geospatial datasets. In previous years, aerial photographs were used as a primary input to map production, but use of satellite data in particular has become the primary input.

38. The current process of extracting information from satellite imagery is operator-assisted and labour intensive. This includes: inspection, analysis, feature extraction (as in identifying roads, rivers or water bodies) and interpretation (which involves deeper knowledge, i.e. the relationship among climate, geology, soil and vegetation and the ecological processes involved).

Challenges and opportunities

39. Two drawbacks of the current methodology and infrastructure are: (1) at times insufficient resolution of satellite imagery for specified functions, and (2) the labour and time required to extract information from the data and integrate with the geographic register, or products. As the demand for more frequent and detailed map products increases, issues such as data storage, life-cycle management, and processing large and varied volumes of satellite images will have to be addressed.

2. Research and development

40. INEGI is investing significant effort on research and development activities related to EO and satellite data in an effort to reduce the barriers to use, and to better exploit the types of data that can be collected in this way:

Big data & machine learning as an enabler of EO use

41. Big data and the latest machine learning techniques may help to overcome processing and time-intensive processes by learning how to properly train the software with the guidance of human experts. The expert will create training databases for the algorithm and will use the same technique to measure the quality of the results.

Agricultural statistics

42. Satellite images ease field work carried out in censuses and surveys related to agriculture. Each surveyor has a mobile device containing satellite images and vector layers that cover their work area; they then use such information to obtain feedback regarding the current limits of the land. The feedback updates the data utilized in later versions of these censuses and surveys. Additionally, the images act as support for orientation while working at the field to agriculture and other sectors of INEGI.

43. Research has been done for identifying the coverage of specific crops in satellite images. Results of these practices are: (1) the construction of a field verified database, (2)

its application for classifying larger areas, and (3) work began on the development of an area sampling frame for agricultural statistics based on satellite images and field data collected from all croplands in 2016.

Geospatial Data Cube of Mexico (information management)

44. INEGI is working jointly with Geoscience Australia in the capacity building for implementing the Mexican Geospatial Data Cube; the cube will feature a nationwide database for archive and recent satellite images, along with the tool for its digital analysis. This database will achieve standard management regardless the source sensor.

45. The development of a more coherent approach to data management of the truly enormous storage requirements of such imagery may propel the processes involved in exploiting EO data. If the data can be managed and accessed more efficiently, they will more likely find use in updating information products, monitoring changes, and generating useful information for surveys and censuses.

46. This exciting Geospatial Data Cube research will help to develop:

- Capacities in: (1) remote sensing, (2) data science, (3) process engineering, and (4) software architecture;
- Infrastructure featuring: (1) efficient access to satellite images, (2) broadband to collect large and frequent data, and (3) big data technology to store and to process satellite information.

C. Statistics Canada

47. Statistics Canada does not conduct the collection, pre-processing, and processing of EO data. Rather, the agency makes use of data that is already prepared, generally by a federal partner such as Agriculture and Agri-Food Canada or Natural Resources Canada. The growing importance of EO data to Statistics Canada has been recently strengthened in the form of a Memorandum of Understanding supporting explicit bi-lateral initiatives with Natural Resources Canada (Rancourt and Shukle, 2017). Statistics Canada's main activities in this area as follows:

1. Input to officially published statistics

Crop Condition Assessment Program[1] - The main output of this long-running program is an interactive web-mapping application that depicts current and historical crop and pasture conditions across Canada. The application is mainly used by Government for policy purposes, by the grain marketing industry, researchers, and the farming industry.

Crop Yield Forecasting[2] - This program uses a methodology for producing crop yields without contacting farm operators. A modelling program based on multivariate regressions uses NDVI from Earth Observation data (AVHRR 1 km), agri-climatic indicators derived from weather station data, and historical results from traditional surveys. These results were first published in 2015, with the traditional survey being replaced in 2016. The method can be reproduced in other locations if the input data exist, providing an

[1] Statistics Canada, 2018, Crop Condition Assessment Program, http://www35.statcan.gc.ca/CCAP/en/index
[2] Statistics Canada, 2017, Model-based Principal Field Crop Estimates, http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5225

opportunity to reduce response burden, survey costs and publication delays. The accuracy of the modelling approach is comparable to the traditional survey.

Environmental Accounts - Satellite earth observation (SEO) and SEO-derived products have been used by the Environment Accounts and Statistics Program (EASP) for at least the last 25 years to produce estimates of land cover.3 In the last few years, new SEO-derived products from federal departments and others have become available, and these products have been used to generate official statistics for land and ecosystem accounts,[4] freshwater ecosystem accounts,5 and forest statistics.6 SEO is particularly well suited to assist in the production of some environmental accounts and statistics as it allows for spatial delineation of different land covers, which is necessary to produce land and ecosystem accounts.

2. Research and development

48. Within the context of a rapidly changing data environment in this field, Statistics Canada is pursuing the following research directions:

Crop area estimation - developing a method to produce crop area estimates to feed to the agriculture survey program. If successful, it will help further reduce response burden for farm operators. This approach uses annual land cover and crop classification from medium- resolution Earth observation data (Landsat and Sentinel), in conjunction with provincial crop insurance data to train the classification and validate/adjust area estimates.

Greenhouse area estimation - This research is developing a method to detect and estimate greenhouse areas using EO and other non-spatial data. It will help reduce response burden for farm operators. This research could benefit others for the detection of other types of features from satellite imagery.

Satellite data for mapping built-up area - This activity will determine if EO data can provide improved delineation of the land use reference layer used for ecosystem accounting. Main aspects include use of different SEO and socio-economic data, development of predictive models, and testing whether extent and intensification of built-up area can adequately detected.

Wetland detection - This activity will determine if EO data can provide improved delineation of wetlands as input to environmental accounts, and assessment of short cycle climate change / water stress. Use in revision of flood-plain and risk identification is another application.

Construction starts/completions - Objective is to explore if satellite data can be used to cover the start of construction for all building types across Canada on a monthly basis for high-growth areas. If successful, the modelled data on capital investment would feed several programs on the economic statistics area, and would help improve coverage of the non-residential sector.

[3] Statistics Canada, 2000, *Human Activity and the Environment 2000*, Cat. No. 11-509-XPE, map 3.2, Land Cover 1992

[4] Statistics Canada, 2016, "The changing landscape of Canadian metropolitan areas," *Human Activity and the Environment*, Cat. No. 16-201-X, https://www150.statcan.gc.ca/n1/pub/16-201-x/2016000/section_3-eng.htm

[5] Statistics Canada, 2017, "Freshwater in Canada," *Human Activity and the Environment*, Cat. No. 16-201-X, https://www150.statcan.gc.ca/n1/pub/16-201-x/2017000/sec-3-eng.htm

[6] Statistics Canada, 2018, "Forests in Canada," *Human Activity and the Environment*, Cat. No. 16-201-X, https://www150.statcan.gc.ca/n1/pub/16-201-x/2018001/list_m-c-eng.htm

D. Austria

1. Input to officially published statistics

49. The use of satellite and earth observation data has not yet found its way into standard techniques for existing official statistics, as Statistics Austria currently benefits from access to high quality administrative data sources of all sorts. Any EO data the agency does use is not explicitly managed as part of a systematic input to the geospatial data framework.

50. However some external products created from satellite and earth observation data are being used for some specific tasks. Examples include dissemination products such as the delineation of settlement areas and localities (GSBPM 5.1 – Integrate data; GSBPM 7.2 – Produce dissemination products).

51. In the area of agricultural statistics, the Farm Structure Survey (FSS) is currently fortunate enough to rely on very detailed administrative data sources stemming from the subsidy system. The areas are available as digitally collected plots, as farmers have to exactly demarcate their crops on a digital farm map. Crop estimates are based on the labour-intensive and voluntary help of harvest consultants estimating the crop yield.

2. Research and development

52. Despite the above for existing statistical programs, significant research effort is under way on how to use satellite and earth observation data in an expanding range of statistical programs:

Using airborne laser data (ALS) to improve the building register

53. Airborne laser scanning data is available as open data and is being used to test the completeness and accuracy of building coordinates, and the buildings' characteristics "built-up area" and "building height" in the buildings and dwellings register. The aim is to improve the quality of these characteristics directly in the buildings and dwellings register. One obvious difficulty is the differing definitions of buildings. From ALS data the outer building envelope of all building-like objects can be discovered, also marking e.g. camping cars as small buildings. Another weakness relates to multi-unit structures such as row houses or blocks of buildings from the ALS data. (GSBPM 3.5 – Test production system)

Satellite data as input to valuation of real property

54. Most of the information for the value is taken directly from the sale contract (address, price, size in m^2 of building and of property, etc.). This is enriched by further information on the building taken from the buildings and dwellings register, the average land price on municipality level as well as average income information and demographic variables of the 50 closest neighbours. Features detectable from satellite and earth observation data could provide additional information for the estimation of a real estate value. Examples include: building types, property sizes of nearby estates, existence of green areas in the neighbourhood and possibly solar panels on buildings in the neighbourhood. So the information needed could help to decide which are poor/run down areas or rich areas. (GSBPM 2.3 – Design collection; GSBPM 3.2 – Build or enhance process components).

Machine learning using EO as input

55. An EU-funded project is currently running at Statistics Austria trying to combine various data sources including satellite and earth observation data to improve the spatial resolution poverty indicators and machine learning algorithms. The idea is to present poverty indicators such as the poverty risk in small scale areas. The use of satellite data is

being evaluated and could provide further interesting inputs. (GSBPM 2.3 – Design collection; GSBPM 3.2 – Build or enhance process components)

3. Summary

56. The use of satellite and earth observation data is growing in importance as support information for some statistical program areas. The most useful application of satellite and earth observation data for agricultural statistics would be to achieve better and automatic estimates for the crop yield being fully aware of the complexity and need for a high degree of specialisation. (GSBPM 4.3 – Run collection). The most obvious field of application would be to achieve a methodology to classify and record land cover and land use and to monitor its change. (GSBPM 3.2 – Build or enhance process components).

57. Challenges remain on how to deal with differing spatial resolutions and time stamps as an input to integration activities. To discover and make out the shape and size of a building, the spatial resolution of satellite data has to be finer, and therefore, Statistics Austria is experimenting with image data and in situ data from aerial surveys, which has a finer resolution.

VI. Impact of crises on the statistical area

58. A detailed treatment of this subject is outside the scope of the review team's work. It would require more in-depth study to realistically describe how financial or political crises might affect the stability of the international public / private infrastructure supporting the constellation of satellites, sensors, computing environment, and finally the agencies and staff that access, interpret, and apply EO data as input to official statistics programs.

59. It can be readily surmised from the above, however, that there is a complex interplay of actors, funding bodies and political contexts supporting the EO infrastructure. Within such a context, it is difficult to say how a particular crisis may exert an impact, except perhaps to say that onset, triggers, and effects would be complex; perhaps rolling-out over a period of months or a few years rather than as a short-cycle.

60. If there are significant concerns about the stability of the international EO infrastructure, then it is suggested that a study be commissioned specifically on this point, and that the research team includes representation of financial and managerial experts in this area.

VII. Main findings, emerging issues and new opportunities for NSOs

61. This chapter consolidates the "bottom line" findings, issues and opportunities identified above to set the current state of practice in EO use in NSOs in clear terms. These points then serve as a framework for further discussion by the NSO community on whether and how to advance the use of EO as input to official statistics.

A. Current uses of EO by NSOs

62. The review revealed significant activity in:

- Agricultural statistics programs such as crop type, area, condition and forecasting;

- Environmental accounts modules related to stock of land by class, forest, water, land-use change;
- Nascent programs around sustainable development indicators such as land-use, climate change, water stress, water quality;
- A key point made by the EO community is that the technology lends itself well to several areas of macro-level statistics related to sustainable development indicators. Several examples are given, particularly by CEOS (2018).

B. NSO-EO research

63. The review indicates that research is proceeding in a few key areas:

- Exciting work by Australia and Mexico on Geospatial Data Cubes is important to watch, because it is attempting to break new ground on two fronts: a) how to process vast amounts of EO data to bring new value, and b) to create a business case that will reduce costs for the storage and use of EO data. Further cost reductions, particularly via a shared infrastructure, may help to expand the number of uses of such data.
- Austria and Canada's work on EO data as input to real-estate property valuation, and in updating of the building register are of interest in order to determine if a compelling business case can be made in these areas relative to current approaches (survey or administrative datasets). Austria's research in the use of airborne LIDAR data is a novel approach in applying surgical precision of accurate sensors at close range, for specific areas of concern, in the update of statistical registers. The results of this research will be eagerly anticipated.
- Several member countries and some international organizations are involved in refining EO methods for use in the area of SDG reporting, which appears to be the domain of most significant potential for NSOs and international organizations. EO data plays to its strength in this context - which is to be able to gather macro-level data for large areas irrespective of ground-based infrastructure, settlement, or population.

C. Issues and their implications

64. NSOs enter EO waters at their own risk! Issues include:

(a) **NSOs lack expertise.** NSOs do not necessarily have the knowledge to readily process, interpret signal, and then integrate EO data into existing programs because this science uses concepts, software, and techniques that are quite different from those needed in survey statistics or even for the usual type of administrative data input. Options are to develop such expertise, or to purchase services. These are significant choices, with impact on budget, recruitment / development of new skillsets, etc.

(b) **EO data are voluminous.** Most NSOs are not experienced with the truly massive volumes of data produced by EO sensors. The IT infrastructure required to house and process such data need to be well-understood in advance. This reality is one of the compelling reasons to considering a centralized infrastructure serving multiple clients, such as is done by national governments, and most private imagery firms selling data.

(c) **EO data alone cannot produce the statistics**. NSO program managers run the risk of assuming too much from EO-related transformation projects. EO data require a lot of processing and calibration against other datasets to produce statistics on a given theme. Frequently, they are combined with other available data-sets, at least initially. This often comes as a surprise to programme managers who must justify larger budgets on what they thought would be an efficiency. Projects to transform statistical production should always proceed with comprehensive study of requirements at the NSO level followed by a review of a few different potential solutions, of which EO is one.

(d) **Institutional Commitment to EO Integration.** Like most true infrastructure, modernizing business processes to incorporate EO requires time, budget, and conviction that it will result in a net improvement. Thorough review, excellent planning, and shrewd implementation will help to ensure success for well-defined applications.

D. Opportunities in EO data

65. If approached shrewdly, there are opportunities in EO data:

1. EO strength is macro-level reporting

66. The reason EO has grown substantially over the decades is because it is extremely useful for monitoring macro-level phenomena such as environmental change, agriculture, forestry, geology, etc. Google Earth has brought EO data into the lives of average citizens across the globe. Within NSOs, its use for environmental accounting, agriculture, forestry, land-use, etc. has been proven. Emphasis in these programs should shift to reducing costs through collaboration and refinement of approaches. As this progresses, it is likely that further improvements in spatio-temporal resolution will continue (e.g. greenhouses / solar panel detection), providing new opportunities for enriching these program areas.

2. SDG reporting is a good application of EO technology

67. The macro-level approach to EO collection and reporting lends itself well to SDG reporting at the national, and international levels, and this appears to be the domain of most significant potential for NSOs and international organizations. It has been reported that of the 30 SDG indicators where EO data has a role to play, only 12 have established methodologies and regular data production by countries. This suggests that there is still significant unrealized potential for EO data to contribute to the global SDG indicator framework. A key opportunity then relates to NSO collaboration on the development of standard approaches to collecting and reporting EO-derived data supporting SDG reporting objectives. Coalescing around SDGs as a conceptual framework will result in more rapid advancement in releasing internationally consistent data as input to discussions on this front.

3. NSOs should collaborate on a generalized approach to EO-data use

68. Owing to its cost and the number of satellites and sensors in orbit, the variety of EO data available is not infinite. In order to meet program needs (on SDGs for example) faster and at less cost, NSOs should explicitly collaborate to develop more standard approaches in how such data should be processed, interpreted, and then integrated. Once the “recipe” is reasonably stable, it can be replicated across NSOs. This approach will effectively address

the barriers to entry of cost and skillset, and provide EO data producers with a more diversified (stable) funding pool, giving them more room to invest in improvements.

4. NSOs should consider the feasibility of a consolidated approach to EO data processing

69. Should it be agreed that meeting the SDG national - international reporting goals is a useful context in which NSOs can collaborate on its use, then a natural extension is to explicitly collaborate with national / international EO organizations to agree on the type of inputs needed, and how they should be processed. This discussion would allow NSOs to speak as a single voice to EO data/service providers such that any data produced could be applied in the same way. The advantage is consistency of inputs of this type to official statistics. Once the production process is stabilized, further rounds of research can be scheduled to determine if substantial improvements can be made.

70. As the state of practice for SDGs matures, the development of a common Analysis Ready Dataset (ARD) of EO data, where time series EO data is corrected to bottom of atmosphere reflectance and is in cartographic geometry and stored in a database, would allow for NSOs to work from a common dataset. The growing data asset would serve as the physical manifestation of the results of the collaboration. Priority topics could be researched, with findings implemented across countries, to help avoid duplication of effort.

VIII. Conclusions and recommendations

71. The use of EO data by NSOs for producing official statistics has a long history, and has **reached early maturity in many countries in the areas of agricultural statistics and environmental accounting**. On the research front, emphasis is being given to matters of data storage, access, and value-added processing (Australia, Mexico), further extensions in agriculture (Canada), and on potential for maintenance of building registers - the built environment (Austria, Canada).

72. At the current time, the international NSO community advances its practice largely by agency-specific interests and capacity, **with results being reported in various fora, but not necessarily to the NSO community**. In parallel, there are several international organizations whose primary focus is EO production, processing, and application, but NSOs are not linked very directly to these communities at the present time. There should be more collaboration in this space such as is occurring in Australia, Mexico, and being re-energized in Canada.

73. The in-depth review makes clear that there are many **significant issues and implications of trying to better exploit EO inputs**. Most importantly, the infrastructure needed to fully support data processing, interpretation and analysis is significant. This comes with implications for budget, as well as new skillsets required to sustain such a capacity. Second, careful analysis of business requirements and possible solutions should be evaluated, as EO data cannot fill all data needs.

74. Despite the sober cautions about EO use, there are **real opportunities for the NSO community, most notably in** the area of national - international **SDG reporting**, where EO plays to its strengths in providing a fairly consistent framework for collecting and reporting on regional-level phenomena, and over time. In terms of official statistics, EO data have proven themselves in areas of agricultural statistics and for environmental accounting. Both of these areas serve as the foundation upon which several SDG indicators are derived, so these programs may require only minor tuning to meet the needs of SDG reporting.

75. However, there are several areas of SDG indicators for which EO-based solutions are promising, but not yet developed. There is clear opportunity here for NSO to bring their skills in data integration to extend the accuracy and value of these data. Doing so will provide a more powerful understanding of the social, economic and environment changes occurring in society at regional, continental, and global scales.

76. To the extent that the international NSO community is being asked to play a role in SDG reporting in the future, and indeed to bring its expertise to the identification of the most meaningful indicators on which to report, and the methods required to report on them, the authors of the in-depth review paper identify the following recommendations for discussion:

A. NSOs should use SDG reporting as a context for our priority EO-based research

77. International attention around SDGs is galvanizing into action. NSOs have an important contribution to make by providing consistent, clear and actionable information as input to decision-making in this area, and EO technology will play an important role in some of the important SDG indicators. NSOs should use the SDG framework to focus research on which indicators can most profitably make use of EO inputs. These decisions can then focus our collective efforts in research, and eventually implementation of successful approaches within NSO program operations.

B. NSOs should collaborate on a generalized approach to EO-data use

78. Following the identification of specific SDG indicators for which EO-inputs are likely to be an important component, the NSO community should consider ways in which to explicitly collaborate on the development of the processes that can leverage EO infrastructure within the Generalized Statistical Business Process Model (GSBPM). The results of the collaboration can be reported and if accepted, be implemented across all agencies. Such an approach will help to ensure steady advance in capacity, in priority areas for reporting, and to yield consistent results where it is applied. This would increase momentum of a truly international community of practice in the subject-matter of EO for NSOs.

C. NSOs should collaborate with EO organizations to consolidate input requirements

79. Once NSOs have centred on how best to apply EO data to measure priority SDG indicators, further advances could be achieved by collaborating with EO national / international organizations to specify the raw sensor data processing, quality control, storage, and access requirements serving as input. This will advance the potential of having “Analysis-Ready” inputs to each agency’s statistical programs. Developments in this direction would go a long way to bringing EO-inputs to a level that is more comparable to many of the other inputs used currently by NSOs, and would further strengthen the culture of international collaboration.

80. In support of this recommendation, the efforts of Australia and Mexico to gain a deep understanding of the type of next generation storage and manipulation environments that can exploit even more information from EO data are encouraged, and the international community should view such work not just as a technical exercise, but also for clues about

the kind of organization required to maintain such an infrastructure for the benefit of NSO clients regionally, or even globally, in the coming years.

IX. References

CEOS (2018a). Committee on Earth Observation Satellites. Retrieved August 22, 2018 from www.ceos.org.

CEOS (2018b). Satellite Earth Observations in Support of the Sustainable Development Goals. Available at: http://eohandbook.com/sdg/files/CEOS_EOHB_2018_SDG.pdf.

EO4SDG (2018). Earth Observations for the Sustainable Development Goals. Retrieved August 21, 2018 from eo4sdg.org.

GEO (2018). Group on Earth Observations. Retrieved August 21, 2018 from earthobservations.org.

Global Strategy to improve Agricultural and Rural Statistics (GSARS). (2017). Handbook on Remote Sensing for Agricultural Statistics. Available at: http://gsars.org/wp-content/uploads/2017/09/GS-REMOTE-SENSING-HANDBOOK-FINAL-04.pdf.

Pringle, M., Schmidt, M., and Tindall, D. (2018). Multi-decade, multi-sensor time-series modelling – based on geostatistical concepts – to predict broad groups of crops. Remote Sensing of Environment, 216 (2018), pp. 183-200. https://doi.org/10.1016/j.rse.2018.06.046.

Rancourt, E. and Shukle, P. (2017). Canada's approach to integrating socio-economic and environmental statistics with geospatial information. Available at: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2017/mtg3/UNECE-GGIM_-_Canada.pdf.

United Nations. (2017). Earth Observations for Official Statistics. Available at: https://unstats.un.org/bigdata/taskteams/satellite/UNGWG_Satellite_Task_Team_Report_WhiteCover.pdf.

UN-GGIM (2018). The United Nations Committee of Experts on Global Geospatial Information Management. Retrieved August 20, 2018 from ggim.un.org.

Acknowledgements and Review Team

This review has benefited from the collaborative efforts of skilled staff in several agencies. These agencies are commended for an effort that will influence and inspire further advancement of common objectives in the development of official statistics in this area. Thank you.

The authors acknowledge the guidance provided by the Steven Vale of the UNECE, who provided quick response to questions and suggestions on emphasis, tone, and on the collaborative approach.

Statistics Canada - Co-Lead

Eric Rancourt (Director General) - Statistical Infrastructure Branch
Tim Werschler (Assistant Director), Nick Lantz - Statistical Registers and Geography Division
Gordon Reichert (Chief), Frédéric Bédard - Agriculture Division

François Soulard (Chief), Mark Henry, Marcelle Grenier – Environment Accounts and Statistics
Elton Cryderman (Chief), Nicolas Martinez – Investment, Science, and Technology Division
Alessandro Alasia (Chief) – Centre for Special Business Projects

Mexico - INEGI - Co-Lead

Abel Coronado - Department of Implementation of Services and Technological Standards
Susana Pérez - Deputy Director General of Economic and Agricultural Censuses.
José Luis Ornelas - Director of Natural Resources.
Alejandro Ruiz - Researcher.

Statistics Austria

Ingrid Kaminger, Sibylle Saul - Registers, Classifications and Geoinformation

Eurostat

Ekkehard Petri, Hannes Reuter (Statistical Officer), Maria-Joao Santos

UNECE

Steven Vale - Senior Statistician and Regional Adviser

* * * * *