# Geoinformation Technology for Analysis and Visualisation of High Spatial Resolution Greenhouse Gas Emissions Data Using a Cloud Platform

Vitalii Kinakh[1][(✉)] , Rostyslav Bun[1,2] , and Olha Danylo[3]

[1] Lviv Polytechnic National University (LPNU), Lviv, Ukraine
kinakh.vitalii@gmail.com
[2] University of Dąbrowa Górnicza, Dąbrowa Górnicza, Poland
[3] International Institute for Applied Systems Analysis, Laxenburg, Austria

**Abstract.** The geoinformation technology for spatial analysis and visualisation of greenhouse gas (GHG) emissions is proposed using Google Earth Engine cloud technology as a key component of interaction with high-resolution spatial data. This technology includes a website for spatial analysis and visualisation of vector data, as well as an interactive site for deeper analysis of raster data on GHG emissions. We use high-resolution vector data of emissions at the level of point, line and areal emission sources, which are converted into raster emission data. Emissions can be analysed within user-created polygons including calculation of the total, specific, maximum or average emission magnitudes. There is also the possibility to fix and select pixels containing a certain interval of emission magnitudes. Using Python's Google Earth Engine module, we have created a website where users can clip raster data from hand-drawn polygons that can be saved on Google Drive. We have also used Python modules (Matplotlib, Pandas, Numpy) for statistical analysis of raster data and histogram construction. Geoinformation technology includes many sectors and categories of human activity included in national inventory reports on GHG emissions, such as those regarding the burning of fossil fuels for power and heat production, within the industrial, agricultural, construction, residential, institutional and waste sectors, as well reports addressing emissions caused by chemical processes. Implementation of the proposed technology is presented using high spatial resolution greenhouse gas emissions data from Poland.

**Keywords:** Geoinformation technology · Greenhouse gas · Emission data · High resolution · Google Earth Engine · Spatial analysis · Raster map · Histogram · Pandas · Matplotlib

## 1 Introduction

In recent decades, scientists have identified global warming as a main cause of climate change and many disastrous changes in different regions of our planet. The temperature increase is caused by anthropogenic greenhouse gas effects, especially those due to the

emissions of carbon dioxide, methane, nitrous oxide and other greenhouse gases (GHG), which are produced as a result of intense human activity. Scientists are now trying to find ways to stabilise and decrease GHG emissions into the atmosphere and increase absorption of these gases.

For scientists, as well as practical specialists, effective methods and tools for the estimation of GHG emissions (also known as GHG inventory) are required to better understand emission processes and structure in order to provide effective control mechanisms necessary to meet emission obligations, emission trade quotes, etc. [1]. Such an inventory will also help to estimate the effectiveness of low carbon technologies, renewable energy sources, etc.

Emission inventories are usually performed at the country or regional level. However, spatial estimates are highly needed for the analysis of geospatial patterns of emissions [2]. Such inventories may be used in atmospheric flow and climate models, and in tools to support administrative decision making at the regional or local level [3, 4]. The main benefit of such an approach is a spatial representation of emissions in the territories where they actually occur. In the most cases, spatial analysis of emissions is carried out using a grid with a determined size of cells. Such analyses include the calculation of emission magnitudes, as well as uncertainties [5], and the results are represented as raster maps. A huge effort has been made to increase the spatial resolution of GHG estimates since a higher resolution better reflects the specifics of territorial emission and absorption processes. Accordingly, the grid cell sizes have decreased from 1° latitude and longitude for global fossil fuel $CO_2$ emissions [6] to 1 km for a global fossil fuel $CO_2$ emissions inventory derived using a point source database and satellite observations of night time lights as proxy data [2].

Nevertheless, other approaches also exist which present GHG emissions at the level of emission sources and are based on vector maps of emissions [7] instead of a grid. Such approaches provide increased resolution of inventory and decreased depth of disaggregation of activity and proxy data, and make possible a spatial inventory for separate categories of human activity, individual GHG, etc. [8].

An increasing number of people are now interested in the reduction of GHG emissions. As a result, there is demand for creating simple and handy solutions for the analysis of emission processes and visualisation of results. The cloud-based platform Google Earth Engine [9] provides many new opportunities for analysis and visualisation of geospatial data, including GHG emissions [10]. At the same time, Python modules (Pandas, Numpy and Matplotlib) provide statistical analysis for creating simple histograms for data distribution, etc.

The aim of this study is to present how cloud platform Google Earth Engine and Python modules can be used to create simple and handy geoinformation technology (GIS) for rapid analysis and visualisation of high spatial resolution GHG emissions data in an easily understandable form.

## 2    Specifics of Spatial Data Used

Spatial vector data on GHG emissions in Poland [11] was used as input data for the geoinformation system developed in this study. These data were calculated for different types of emission sources depending on the specifics of each category of human activity. Objects with significant emissions (e.g. power plants, cement plants, refineries, etc.) are considered as point-type emission sources and compared with the surrounding area [12, 13]. Lengthy objects with small widths that cause GHG emissions (e.g. road segments) are considered as line-type emission sources [14]. Objects that emit GHG or absorb the carbon dioxide from some area (e.g. settlements in the residential sector [15], arable lands in the agricultural sector [16], etc.) are considered as area-type source/sink emission sources.

The above mentioned vector data include emissions of carbon dioxide, methane, nitrous oxide and other GHG, as well as the total emission of all gases in $CO_2$-equivalent. The spatial resolution of the emissions data from point- and line-type sources is quite high (i.e. several metres) as all point sources were inspected visually using Google Earth, and line-type sources were built on the basis of Open Street Maps [17]. In contrast, the spatial resolution of area-type sources is lower because they were created on the basis of the Corine Land Cover Maps [18] with 100 m spatial resolution. Therefore, the spatial resolution of these data does not generally exceed 100 m.

## 3    GIS Technology

### 3.1    Structure of Technology

The main components of the proposed GIS technology for spatial analysis and visualisation of high-resolution GHG emissions data are:

- operations with vector data on GHG emissions including website for spatial analysis and visualisation of vector data on GHG emissions at the emission source level;
- conversion of vector data to raster including grid creation and calculation of gridded emissions;
- operations with raster data on GHG emissions including: spatial analysis and visualisation of high-resolution raster data on GHG emissions at grid level using the interactive website; band choice for deep analysis of GHG emissions; polygon creation for deeper analysis of GHG emissions; calculation of the main parameters of emission; fixing interval (i.e. min, max) and selecting pixel values within this interval.

The main structure of this technology is shown in Fig. 1. We can see that all of the above mentioned components are connected with the Google Earth Engine cloud platform as a core of the developed geoinformation system.
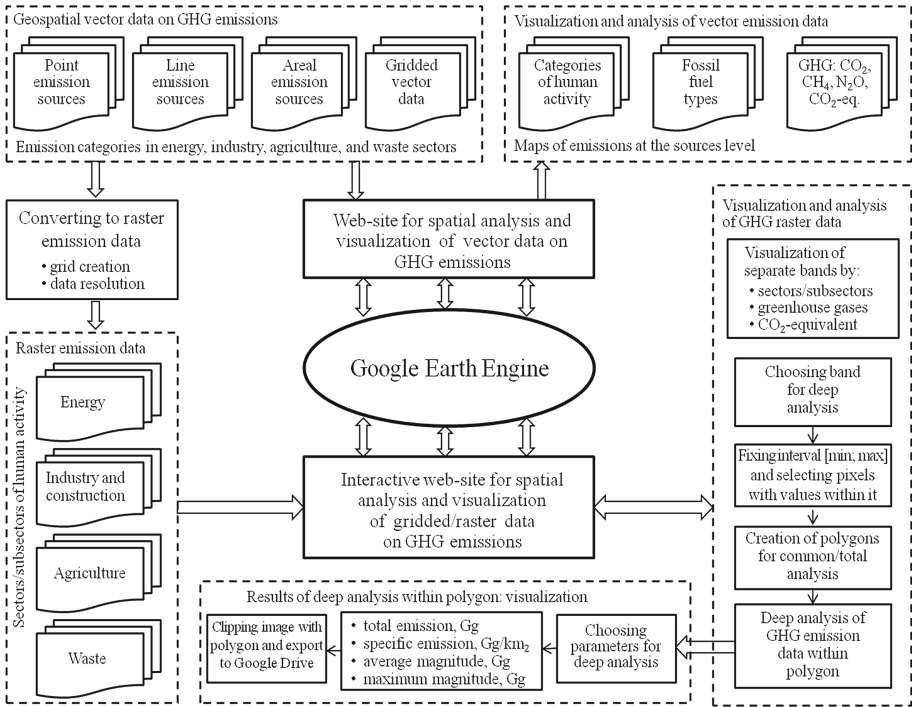
**Fig. 1.** Structure of the GIS technology for spatial analysis and visualisation of high-resolution GHG emissions data

## 3.2 Main Components

The website for spatial analysis and visualisation of vector data on GHG emissions at the emission source level was carried out using Google Maps JavaScript API. The vector spatial data were uploaded onto Google Fusion Tables and made accessible from Google Earth Engine. These data are based on a high-resolution spatial inventory [11–16] and include emission magnitudes of separate GHG from point-, line- and area-type emission sources, as well as total emission of all gases in $CO_2$-equivalent. Additionally, all spatial data on emissions were aggregated into cells of a fixed size. A $2 \times 2$ km grid was used for the emissions data from Poland. Cells are further divided into smaller vector objects if they are crossed by administrative boundaries of a province, district or municipality, which helps to maintain administrative assignment of each elementary object and provides the opportunity to sum GHG emissions within the bounds of each administrative object without loss of precision.

The main raster parameters, which are related to spatial resolution and map extent, are set during the conversion of vector data (usually in tab MapInfo or ESRI shape file formats) into raster/gridded data (in tiff file format). High spatial resolution GHG emissions data from different sectors and categories of human activity are organised as bands of raster data. The bands contain information regarding the emission of different GHG

(carbon dioxide, methane and nitrous oxide), as well as the total emission of all gases taking the global warming potentials of separate gases into account.

The interactive website for spatial analysis and visualisation of high-resolution raster data on GHG emissions at the grid cells level was also carried out using Google Maps JavaScript API. The additional functionality of this site provides the possibility to choose the corresponding raster data band, create any arbitrary polygon and analyse the main parameters of emission processes within its bounds. The user can chose an operation such as calculation of the total, specific, maximum or average emission magnitude. Users can also fix minimum and maximum values of some interval and then select pixels for which the magnitude is within this interval. The raster data include carbon dioxide, methane and nitrous oxide emissions, as well total emissions in $CO_2$-equivalent in the energy sector, separately for electricity generation and heat production, extraction and processing of fossil fuels, petroleum refinement, transport, as well as for the manufacturing, industry, construction, commercial, residential, agricultural and waste sectors, among others.

The second interactive website for spatial analysis and visualisation of raster data was developed using Google Earth Engine Python API. This tool provides additional functionality such that users can draw a polygon, clip all cells within the polygon as a separate raster image and export it to Google Drive. This functionality is a great opportunity for users who do not have powerful computers as all operations are processed in the cloud.

There are several positive features of using such a tool for spatial analysis of GHG emissions. First, as the tool is based on Google Earth Engine, there are no strong requirements of the user's computer for interacting with it, aside from Internet access and the availability of a browser with JavaScript support. Second, all data are stored on the server such that users do not need to download and pre-process the data. Finally, the specified emissions in a selected sector can be calculated for the area of interest without handling the entire dataset.

In order to calculate total emissions, the analytical results at the level of separate emission sources are aggregated to vector data in a regular $2 \times 2$ km grid. As a result of this operation, emissions from point-, line- and area-type sources that are completely or partially within these cells are summed [11]. The easy operation of these data and quick implementation of analysis of the main parameters of the emission processes within the user-determined territories demonstrates the need to create corresponding geoinformation technology using Google Earth Engine.

## 4   Practical Implementation of Gridded Emissions

Practical implementation of the geoinformation technology for spatial analysis and visualisation of high-resolution GHG emissions data is presented in Fig. 2 using the example of total GHG emissions in the carbon dioxide equivalent from the residential sector in Poland. The red-yellow colour scheme was selected for better visualisation of the results.
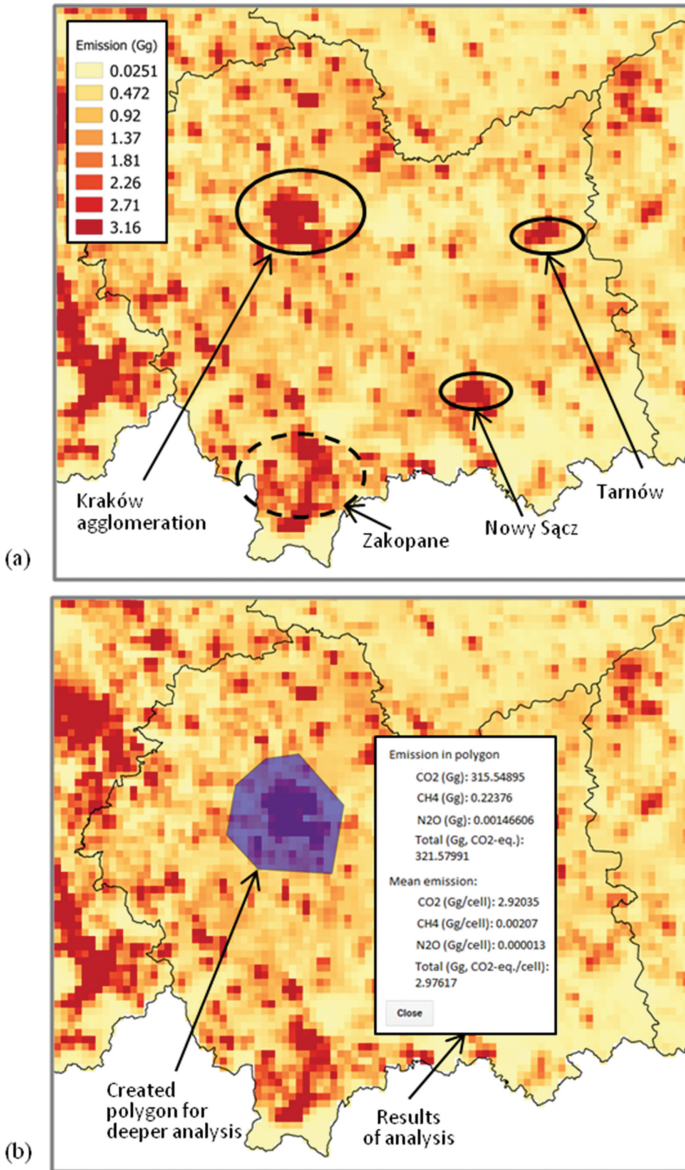
**Fig. 2.** Example of geospatial analysis of total GHG emissions in the residential sector in the Lesser Poland province: (a) initial raster data and (b) polygon created with visual results of a deep analysis (2010, 2 × 2 km grid, $CO_2$-equivalent, Gg)

These GHG emissions data were obtained at the level of settlements as area-type emission sources. For this purpose, the statistical data on fuel consumption at the municipal level were disaggregated into the level of settlements using proxy data including: maps of heating degree days and population density in Poland; data on access to energy

sources; living area data; percentage of living area equipped with central heating and hot water supply and data regarding the amount of heat energy provided to households. A full description of this approach for high-resolution spatial inventory of GHG in the residential sector is presented in [15]. Following this approach, data on the consumption of coal, natural gas and biomass, as well as the corresponding emission factors, emissions of carbon dioxide, methane, and nitrous oxide at the level of settlements and total $CO_2$-equivalent emissions were computed [19].

These vector data on GHG emissions in Poland were aggregated into $2 \times 2$ grid cells as vector map elements and then converted into a similar raster grid in tiff format. Implementation of these spatial data into the developed geoinformation system is illustrated in Fig. 2 using the Lesser Poland province as an example. The upper part of Fig. 2 demonstrates gridded emissions in the residential sector. One can see the highest emissions in Kraków agglomeration and emissions of the largest towns of this region including Tarnów and Nowy Sącz. In the bottom of Fig. 2, one can see emissions from the neighbourhood of Zakopane. Although this territory is a resort area, it is densely populated and wood is primarily used as fuel within the residential sector, which has a higher emission coefficient compared with other fossil fuel energy sources.

In Fig. 2b, one can see how the interactive website works. In particular, a user-formed polygon around the Kraków agglomeration is shown with a box of results from additional deep analysis of emission processes. Generalised parameters (e.g. mean and total emission values within the polygon) are indicated within this box.

In order to create the website for analysis and extraction of images/maps, we used Google Earth Engine Python's API, Jinja2 HTML template (i.e. website templates) and Google Drive API for saving images. Users can download images clipped within the hand-drawn polygon for further analysis. This functionality is very helpful for inexperienced users who do not have powerful computers as such operations using a desktop GIS involve significant computational costs and all operations on our website are processed in the cloud.

## 5 Statistical Analysis of Gridded Emissions

The geoinformation technology for spatial analysis of high-resolution GHG emissions data developed in this study provides the possibility to calculate some additional parameters of the raster data. For statistical analysis of emission magnitudes in pixels and the creation of histograms and plots, we used the Python modules Matplotlib, Pandas and Numpy [20]. We used raster maps in ESRI shape file format of high-resolution GHG emissions data from Poland as input data [12–16]. Then we selected pixels with non-zero emission magnitudes using the Pandas module. Due to the specifics of emission processes in some categories of human activity (e.g. electricity generation, petroleum refinement, metallurgy, etc.), there are a significant number of cells with a zero emission magnitude. Using this dataset and the Matplotlib module, we created histograms of emission magnitudes in $CO_2$-equivalent for all sectors/subsectors covered by the high-resolution spatial inventory of GHG. Plots and maps can be manipulated in two different ways using Matplotlib. The first is an interactive way where one can zoom into a plot,

view a chart from different angles and view specific emission values or regions. The second way is to save and work with the plot as a picture.

We show histogram examples in Fig. 3 for GHG emission magnitudes in Poland per each $2 \times 2$ km cell created for:

– fossil fuels (i.e. liquid, solid and gaseous fuels) in the residential sector;
– wood/biomass in the residential sector;
– other categories including agriculture, forestry, fishery etc.

We show similar histograms in Fig. 4 for GHG emission magnitudes per cell from the use of fossil fuels in the:

– manufacturing industry and construction sectors;
– transport sector;
– energy sector (total emissions).

Both figures include calculated histograms for total emissions of all GHG (carbon dioxide, methane and nitrous oxide) in $CO_2$-equivalent terms. For better visualisation of the results along the ordinate axis, the occupied area is displayed instead of the number of pixels.

The histogram for each sector consists of two parts. The upper part demonstrates the distribution of data from minimum to mean value. This is the main part of the GHG emissions data and reflects emissions from the majority of cells. The lower part of each histogram illustrates the full data distribution (i.e. from minimum to maximum). The majority of the data is clearly within the first part. This feature is characteristic for most sectors including residential, agriculture/forestry/fishery, manufacturing industry, construction and transportation.

Emissions within the pixels from the residential sector are relatively small, both from the burning of fossil fuels and from the use of wood/biomass (Fig. 3). For fossil fuels, a large area (124,000 km$^2$) is occupied by pixels with emissions less than 150 Mg$_{CO2\text{-}eq}$. Similarly, for the use of wood, an area of 127,000 km$^2$ is occupied by pixels with emissions less than 80 Mg$_{CO2\text{-}eq}$. Fossil fuel emissions within the agriculture/forestry/fishery sector are insignificant such that more than 89,000 km$^2$ occupy pixels with emissions less than 20 Mg$_{CO2\text{-}eq}$.

For non-zero pixels of the manufacturing industry and construction sector in Poland, the mean magnitude is higher (4.34 Gg$_{CO2\text{-}eq.}$) (Fig. 4), which is influenced by high emissions point sources (e.g. metallurgical and chemical plants, etc.) even though the total area occupied by these pixels is relatively small. The most 'uniform' emissions distribution is in the transport sector, for which the mean value for non-zero pixels is 654 Mg$_{CO2\text{-}eq}$.

As can be seen in Fig. 4c, the distribution of GHG total emissions in the energy sector in Poland is close to lognormal. The magnitudes of most pixels with a total area larger than 25,000 km are within the range of 279–419 Mg$_{CO2\text{-}eq}$. The tail of this distribution, however, reaches a maximum value of 26,000 Gg$_{CO2\text{-}eq}$, as can been seen in bottom part of the histogram in Fig. 4c. The highest emissions are caused by power plants as emission point sources although there are only a few such pixels, whereas there are 80 power plants in Poland that use fossil fuels with power greater than 20 MW.
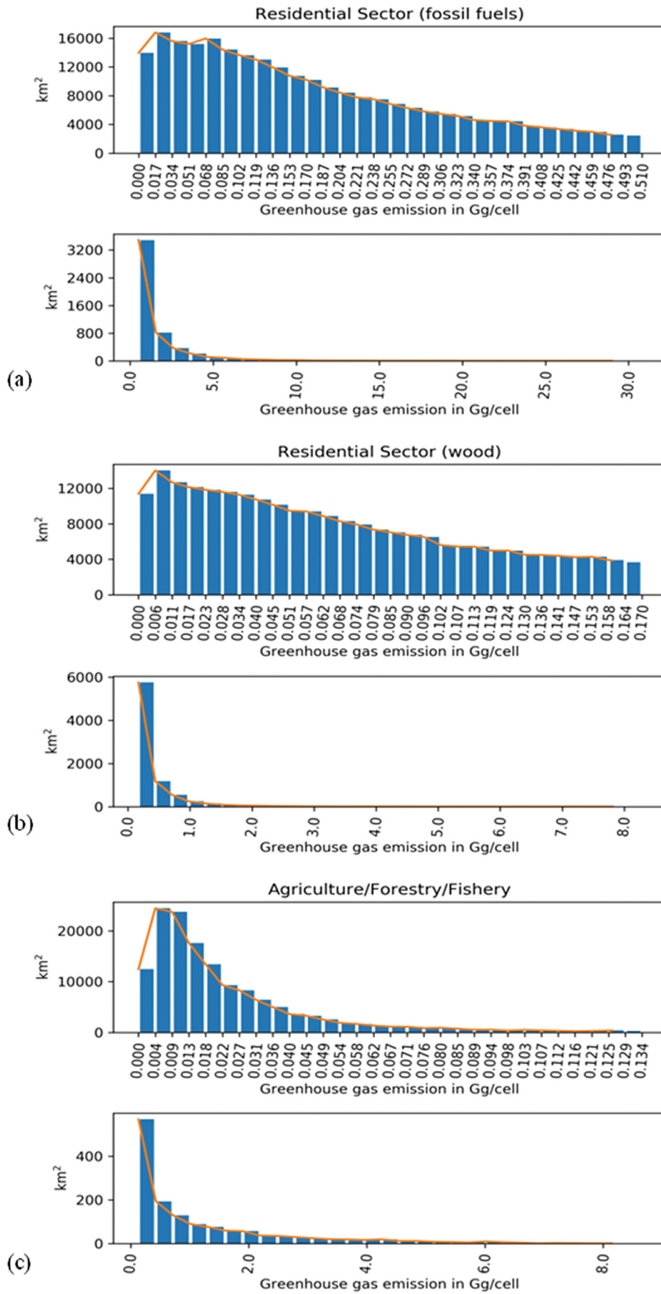
**Fig. 3.** Histograms of emission magnitudes per cell in the residential sector due to (a) fossil fuels, (b) wood and (c) in other categories such as agriculture, forestry etc. (Poland, 2010, $2 \times 2$ km grid, $CO_2$-equivalent, Gg/cell)
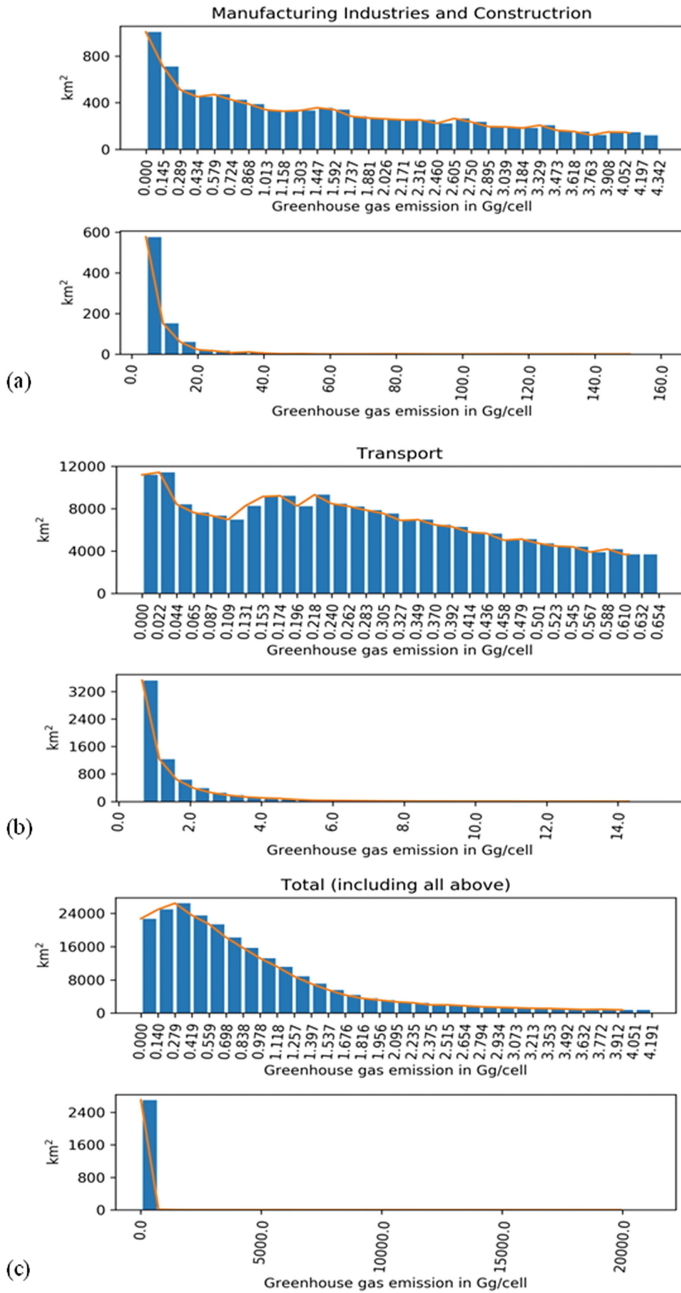
**Fig. 4.** Histograms of emission magnitudes per cell in the (a) manufacturing industry and construction sector, (b) transport sector and (c) energy sector (total emissions) including all subsectors (Poland, 2010, $2 \times 2$ km grid, $CO_2$-equivalent, Gg/cell)

The presented data characterise the uneven distribution of emissions in the investigated region, as well as the expediency of such spatial inventories of GHG and potential of the created geoinformation technology for spatial analysis. In general, all websites and tools developed in this study can be used for spatial analysis of any raster map/image available in Google Earth Engine, not only for GHG data.

## 6 Conclusions

The presented geoinformation technology for spatial analysis and visualisation of high-resolution GHG emissions data uses the cloud-based platform Google Earth Engine. Presented histograms and charts were created using Python modules Matplotlib, Pandas and Numpy. Vector maps of emissions at the level of point-, line- and area-type sources are used as input data. The analytical results are represented as raster emission maps. Implementation of the developed technology is demonstrated using high spatial resolution GHG emissions data from Poland as a case study. This analysis includes emissions in all main sectors and categories of human activity covered by the National Inventory Reports [21] submitted to the United Nations Framework Convention on Climate Change.

The use of vector data on GHG emissions in the created technology yields better results in terms of uncertainty as well as usability. The methodology allows utilisation of high-resolution input vector maps of point-, line-, and area-type emission sources, and maintains high resolution in the results, which are independent of the grid size, overlapping grids, etc. Consequently, it excludes the source location uncertainty as a component of the total uncertainty. The use of raster data on GHG emissions in this geoinformation technology provides the potential to perform effective spatial analysis at the regional and national scales, as well as deeper emissions analyses within user-created polygons, etc.

The Google Earth Engine provides an opportunity to perform spatial analysis of emission processes independent of computer characteristics and user experience with regards to GHG emission inventories. The created geoinformation technology also allows deeper analysis of generalised parameters of emission processes within user-formed polygons with calculation of the total or specific, as well as maximum or average emission magnitudes.

## References

1. Ometto, J.P., Bun, R., Jonas, M., Nahorski, Z. (eds.): Uncertainties in Greenhouse Gas Inventories: Expanding Our Perspective. Springer, Cham (2015)
2. Oda, T., Maksyutov, S.: A very high-resolution (1 km × 1 km) global fossil fuel $CO_2$ emission inventory derived using a point source database and satellite observations of nighttime lights. Chem. Phys. **11**, 543–556 (2011)
3. Déqué, M., Somot, S., Sanchez-Gomez, E., Goodess, C.M., Jacob, D., Lenderink, G., Christensen, O.B.: The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Clim. Dyn. **38**(5), 951–964 (2012)

4. Neale, R.B., Richter, J., Park, S., Lauritzen, P.H., Vavrus, S.J., Rasch, P.J., Minghua, Z.: The mean climate of the community atmosphere model (CAM4) in forced SST and fully coupled experiments. J. Clim. **26**, 5150–5168 (2013)

5. Hogue, S., Marland, E., Andres, R.J., Marland, G., Woodard, D.: Uncertainty in gridded $CO_2$ emissions estimates. Earth's Future **4**, 225–239 (2016)

6. Andres, R.J., Marland, G., Fung, I., Matthews, E.: A $1° \times 1°$ distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. Global Biogeochem. Cycles **10**(3), 419–429 (1996)

7. Bun, R., Gusti, M., Kujii, L., Tokar, O., Tsybrivskyy, Y., Bun, A.: Spatial GHG inventory: analysis of uncertainty sources. A case study for Ukraine. Water Air Soil Pollut. Focus **7**(4–5), 483–494 (2010)

8. Boychuk, K., Bun, R.: Regional spatial inventories (cadastres) of GHG emissions in the energy sector: accounting for uncertainty. Clim. Change **124**, 561–574 (2014)

9. Google Earth Engine: A planetary-scale platform for Earth science data and analysis. https://earthengine.google.com. Accessed 5 July 2017

10. Lemoine, G., Léo, O.: Crop mapping applications at scale: using Google Earth Engine to enable global crop area and status monitoring using free and open data sources. In: Remote Sensing: Understanding the Earth for a Safer World, IGARSS 2015, Milan, pp. 1496–1499 (2015)

11. Bun, R., Nahorski, Z., Horabik-Pyzel, J., Danylo, O., Charkovska, N., Topylko, P., Halushchak, M., Lesiv, M., Striamets, O.: High resolution spatial inventory of GHG emissions from stationary and mobile sources in Poland: summarized results and uncertainty analysis. In: Proceedings of the 4th International Workshop on Uncertainty in Atmospheric Emissions, Kraków, Poland, 7–9 October 2015, pp. 41–48. SRI PAS, Warsaw (2015)

12. Topylko, P., Halushchak, M., Bun, R., Oda, O., Lesiv, M., Danylo, O.: Spatial greenhouse gas (GHG) inventory and uncertainty analysis: a case study for electricity generation in Poland and Ukraine. In: Proceedings of the 4th International Workshop on Uncertainty in Atmospheric Emissions, Kraków, Poland, 7–9 October 2015, pp. 49–56. SRI PAS, Warsaw (2015)

13. Charkovska, N., Halushchak, M., Bun, R., Jonas, M.: Uncertainty analysis of GHG spatial inventory from the industrial activity: a case study for Poland. In: Proceedings of the 4th International Workshop on Uncertainty in Atmospheric Emissions, Kraków, Poland, 7–9 October 2015, pp. 57–63. SRI PAS, Warsaw (2015)

14. Boychuk, P., Nahorski, Z., Boychuk, K., Horabik, J.: Spatial analysis of greenhouse gas emissions in road transport of Poland. Econtechmod **1**(4), 9–15 (2012)

15. Danylo, O., Bun, R., See, L., Topylko, P., Xu, X., Charkovska, N., Tymków, P.: Accounting uncertainty for spatial modeling of greenhouse gas emissions in the residential sector: fuel combustion and heat production. In: Proceedings of the 4th International Workshop on Uncertainty in Atmospheric Emissions, Kraków, Poland, 7–9 October 2015, pp. 193–200. SRI PAS, Warsaw (2015)

16. Charkovska, N., Bun, R., Danylo, O., Horabik-Pyzel, J., Jonas, M.: Spatial GHG inventory in the Agriculture sector and uncertainty analysis: a case study for Poland. In: Proceedings of the 4th International Workshop on Uncertainty in Atmospheric Emissions, Kraków, Poland, 7–9 October 2015, pp. 16–24. SRI PAS, Warsaw (2015)

17. Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (eds.): OpenStreetMap in GIScience - Experiences, Research, and Applications. Springer, Cham (2015)

18. Corine Land Cover data (2016). http://www.eea.europa.eu/. Accessed 20 June 2017

19. Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K. (eds.): IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, IPCC (2006)
20. McKinney, W.: Python for Data Analysis, pp. 120–256. O'Reilly Media, Sebastopol (2012)
21. Poland's National Inventory Report 2012, KOBIZE, Warsaw (2012). http://unfccc.int/national_reports. Accessed 1 July 2017