

12

DISAGGREGATION PROBLEMS USING DATA DERIVED FROM POLISH AIR POLLUTANT EMISSION MANAGEMENT SYSTEM

12.1 INTRODUCTION

The establishment of the NED database at the end of 2010 was strictly associated with the development of the Polish national law on air pollutant emission management [4]. Before, the necessary data has to be collected only for purposes of centralized country-level air emission management in distributed system. Establishment of the NED resulted also with simplification of administrative procedures on fulfilling selected international obligations, described widely in [2] or [17].

Data collected in the NED database is mainly focused on matters of country-level air pollutant emissions management. Structure of arrangement of the data is described in details in adequate national Polish regulation [5]. The simplified structure of the data collection is presented in [21]. There are various pieces of information gathered i.e.: emission (annual amount of air pollutants or greenhouse gases released to the atmosphere), geographical coordinates of emitters (as a part of installation), information on technology and air pollutant control devices, fuel used and other.

Pieces of enumerated data, is being submitted to the National Centre for Emission Management as an annual obligation and is partly applied also for purposes of fulfilling of international obligations (national air pollutant emission inventory [10]).

The application of data provided by the NED database is connected with using selected methods of data integration and correction. These types of operations are frequently applied especially during aggregation of the emission data to obtain the value of national annual emission of air pollutant from particular kind of anthropogenic activity.

12.2 MATERIALS AND METHODS

Bottom-up and top-down approach

Simplified structure of reporting data is presented in the figure below (Fig. 12.1). Particular installation (fuel combustion installation or other kind, defined as significant, in the context of emitted air pollutants [5]) is basic data structure in the NED database. Installation is combined from emission sources (e.g. boiler, furnace, hearth) joint with conduits (not distinguished in the figure above). Installation and emission source are

also joint (using conduits) with the emitter (primarily stack, or another place of release of air pollutants to the atmosphere). Geographical coordinates (longitude and latitude) are assigned to the point of releasing of the air pollutants to the atmosphere (emitter).

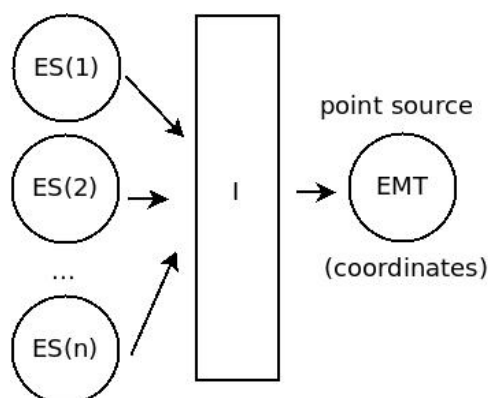


Fig. 12.1 Simplified structure of objects in the NED database (own analysis)
ES(1), ES(2), . . . , ES(n): emission sources; I: installation; EMT: emitter

Presented kind of arrangement system (primarily point sources) is commonly known as the 'bottom-up' approach, considering primarily the point source of emission as a basic activity. The opposite, named 'top-down' – basing mainly on estimations using combination of the regional (national) data derived from statistical surveys (activities of emission sources) and typical, average values of emission of particular pollutant corresponding to the unit of product – emission factors (described in details in [24]).

Data derived from the NED can be classified as the area between 'bottom-up' and 'top-down'. Selected categories (e.g. combustion of fuels in public utility plants) are completely corresponding to the official statistics provided by the Polish Central Statistical Office. For categories related to the surface activity (such as emission of air pollutants from agriculture – especially animal breeding or fertilizers application) the compliance is significantly lower due to various discrepancies, data inconsistency and uncertainties associated with emission estimation.

For practical applications, the analysis using only one of enumerated approaches ('bottom-up' or 'top-down' only) is carrying out very rarely. Commonly during emission inventory compilation the combination of those two approaches is simultaneously used.

Combustion of fuels in stationary sources

Emissions of air pollutants from combustion of fuels in stationary sources are included in the national air pollutant emission inventory [10]. Scheme presenting compilation of 'bottom-up' also 'top-down' approaches is shown in the figure below (Fig. 12.2). Emissions sources described as PP and IPP are derived from the NED database and split between main Polish administrative units (provinces), (area of Poland is split between 16 main administrative units (provinces, voivodships) corresponding to NUTS 2 statistical classification provided by EUROSTAT). The rest of categories (CHP, HP, REF, COKE, MIN) are currently estimated with using top-down approach [10].

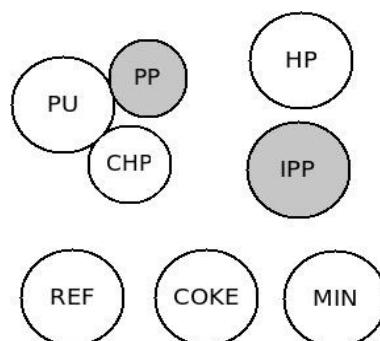


Fig. 12.2 Compilation of bottom-up and top down approaches: air pollutant emissions from the Polish energy sector (combustion plants)

Bottom-up approach marked with grey color.

PU: public utility plants, PP: power plants, CHP: combined heat and power,

HP: heating plants, IPP: industrial power and CHP plants,

REF: petroleum refining plants, COKE: solid fuel processing plants (coke production plants),

MIN: power and CHP plants in coal and lignite mines.

However the category including PP and IPP plants is able to split between Polish provinces manually ('bottom-up' approach), for obtaining the split of whole energy sector the data disaggregation is needed.

Disaggregation and spatial allocation

Disaggregation of air pollutant emissions, known also as downscaling process in this context, is widely described in [1, 3, 6, 9, 12, 13, 16, 18]. In majority of cases is basing on the choice of kind of representative value (a type of disaggregation factor, also combination of several factors), named surrogate. The assumptions are that the surrogate should be transparent for particular sector, corresponding to emission of air pollutant from particular sector (or group of emission sources).

Frequently, as the surrogate is being assumed value directly related to the estimated emission, it can be amount of goods produced (e.g. energy, heat, gasoline or another), fuel consumption, employment in particular sector or also economical factor typical for particular region (e.g. GDP per capita in particular province). Specific kinds of surrogates can be also assumed basing on emission data derived from the data collection such as the NED database (this particular type of splitting factor is known as 'spatial surrogate'). Structure of derivation of surrogates from the NED database is presented in the figure below (Fig. 12.3).

Disaggregation of top-down estimation (split into administrative units or smaller areas, e.g. grids used for purposes of modeling) is connected with problem of choice of proper spatial allocator for emission. Indirect allocator could be established by the value representing the unique features of analyzed area. In case of emission inventory compilation the spatial allocator can be also a result of modeling. Surrogates are not only way of disaggregation. More and more frequent way of splitting emission data are use of spatially oriented statistical methods [8, 14, 7]. The advantage of application spatial statistics in comparison to using surrogates is possibility of work in both scale approach: regional [7] and local [22].

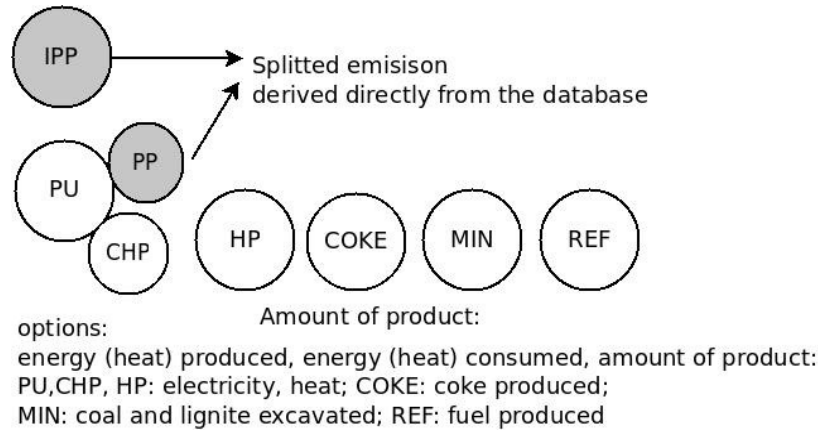


Fig. 12.3 Derivation of surrogates using data derived from the NED database

The one of the most popular way applied for purposes of spatial analysis is kriging. This kind of spatially oriented statistical estimation (geostatistical) is described in details in [7, 20, 23] and many other. Kriging (as a class of algorithms) is a base for another geostatistical estimations and is also described with the acronym 'BLUE' (Best Linear Unbiased Estimator) [23]. The idea of kriging is to estimate the value of continuous parameter z in the neighborhood of the point x_0 with using of linear weighted combination of values of the parameter z from the neighborhood of x_0 using random functions and regionalized variables theory.

Kriging algorithm optimizing the weights of values 'in the neighborhood' to obtain the error of mean value equal to 0 and also minimize the variance of error in that model. The basic procedure of kriging estimation is described in equations below.

$$\hat{Z}(x_0) = \sum_{i=1}^n w_i Z(x_i) \quad (12.1)$$

$$R(x_0) = \hat{Z}(x_0) - Z(x_0) \quad (12.2)$$

$$R(x_0) = \sum_{i=1}^n w_i Z(x_i) - Z(x_0) \quad (12.3)$$

$$E[R(x_0)] = E[\sum_{i=1}^n w_i Z(x_i) - Z(x_0)] = \sum_{i=1}^n w_i E[Z(x_i)] - E(Z(x_0)) \quad (12.4)$$

$$E[R(x_0)] = \sum_{i=1}^n w_i E(Z) - E(Z) \quad (12.5)$$

$$E[R(x_0)] = (\sum_{i=1}^n w_i - 1)E(Z) = 0 \quad (12.6)$$

$$\sum_{i=1}^n w_i = 1 \quad (12.7)$$

Where denoted:

- (12.1) - linear, weighted average of the values 'in the neighborhood',
- (12.2) - residual value: estimation error,
- (12.3) - residual value as combination of $n+1$ random variables,
- (12.4), (12.5) and (12.6) - expected value of estimation error,
- (12.6) - assumed stationarity of random values,
- (12.7) - assumed no bias for estimation.

Data analysis

Derived data on emission of NO_x and SO_x from stationary combustion sources located in Silesian Province (point source data submitted for 2011). Applied procedure of data analysis is shown in the figure below. The data was corrected before, accordingly to the procedure described in [22]. Before application of the procedure described above, the spatial continuity is analyzed using variogram analysis (described in details below) (Fig. 12.4).

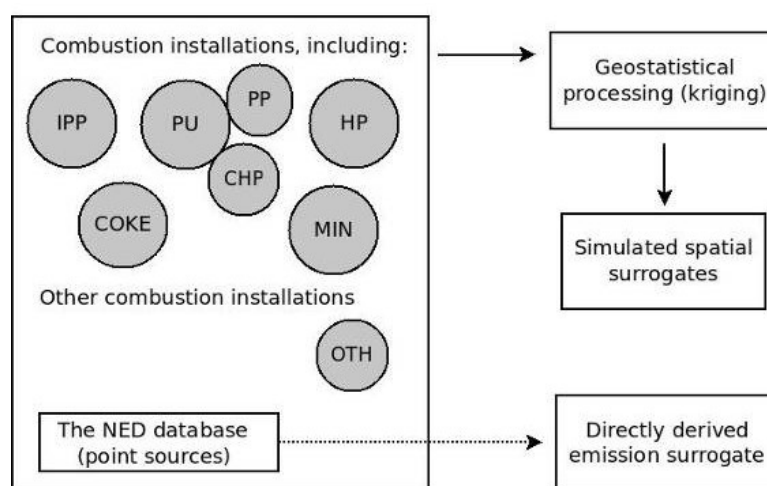


Fig. 12.4 Applied procedure of data analysis

For purposes of analysis there was applied spatial continuity analysis using variograms. Variography described in details in e.g. [20, 23] was applied to originally submitted data to the NED database (as above, considered as point source emissions). Theoretical semivariograms for purposes of kriging estimation (for analyzed pollutants) were built basing on empirical (generated) variograms (variogram clouds). Manual fitting was applied. Kriging simulation was applied to the anisotropic semivariograms, which were chosen due to analysis of stationary sources.

Ordinary kriging estimation was applied in areas of size 500×500 [m²] using geostatistical R package 'geoR' [11, 19]. Results of conducted analysis are presented in the next section.

12.3 RESULTS AND DISCUSSION

Empirical semivariograms

Basing on generated variogram clouds (presented in the figure below) there were assumed theoretical semivariograms. Parameters of manual fitting are enumerated below:

NO_x : sill = 8 [$\log_{10}(\text{kg})^2$], range = 1000 [m],
 practical range = 2996 [m], nugget = 2.6 [$\log_{10}(\text{kg})^2$],
 SO_x : sill = 9 [$\log_{10}(\text{kg})^2$], range = 1500 [m],
 practical range = 4494 [m], nugget = 2.7 [$\log_{10}(\text{kg})^2$].

For both pollutants was assumed exponential model of semivariogram. Generated variogram clouds with fitted formulas (basing on parameters included above) are presented in the figure below (Fig. 12.5). Semivariograms presented above are typical for large datasets with discontinuous values. Shape of variogram clouds suggests that should be used auxiliary statistical tool or supportive data transformation more advanced than logarithmic [15].

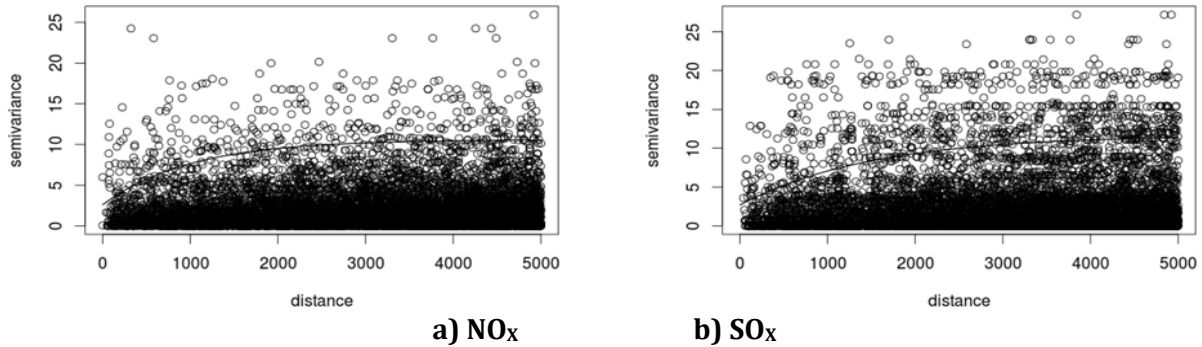


Fig. 12.5 Variogram clouds and assumed theoretical semivariograms

distance = [m], semivariance = $[\log_{10}(\text{kg})^2]$

Ordinary kriging

Simulated values of annual NO_x and SO_x emission from stationary combustion sources (included standard error values) are shown in figures below (Fig. 12.6). In that particular case, the kriging process is both: an estimation (estimates spatial distribution of emission) and also simulation (distribution is simulated due to application random variable theory and functions of co regionalization). Presented values for whole area of Silesian Province are limited only for stationary combustion sources. It can be observed not very strong correlation between results for NO_x and SO_x .

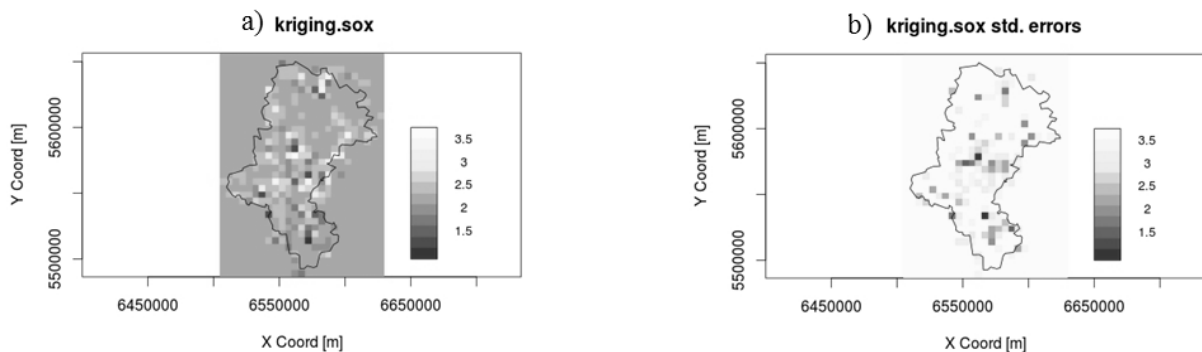


Fig. 12.6 NO_x

Values marked red (kriging) are located in places with the lowest values simulated (Fig. 12.7). In opposite – the highest values of standard error of estimation are located comparably, in places of the lowest input values of emission. This fact indirectly confirms hypothesis on occurrence of discontinuous areas. Anisotropic semivariograms were applied due to constant spatial distribution of analyzed objects (fuel combustion installations).

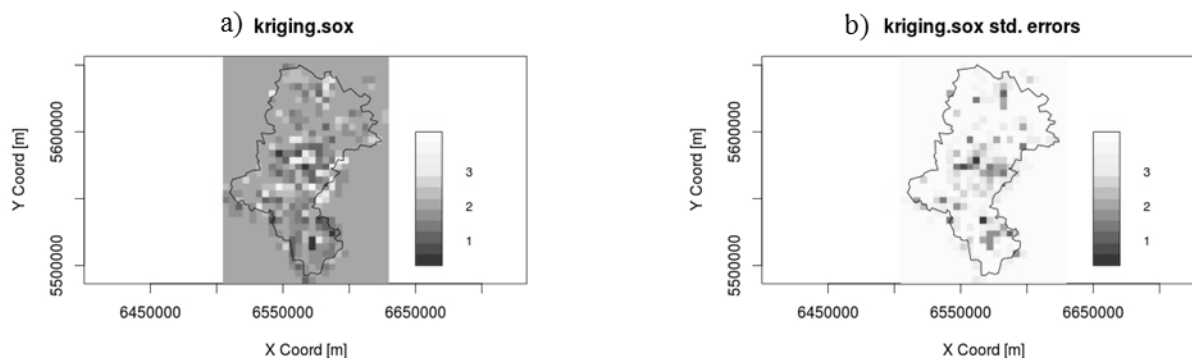


Fig. 12.7 SO_x

CONCLUSIONS

Geostatistical analysis and modeling is becoming more and more popular device applied for spatial data, especially in cases of analysis of discontinuous areas. Specific case of Polish conditions, where the significant part of air pollution is still generated by combustion processes, causes that analysis of environmental effects of the power and energy sector is still current topic.

Application of elements of geostatistical analysis can be potentially high-effective auxiliary tool in emission inventory, especially in case of analysis emissions from stationary sources also with selected downscaling problems, such as disaggregation. Results of that kind of analysis can simplify obtaining spatial surrogates sufficient for dealing with current problems of emission management, especially downscaling the emission data. Geostatistical methods, such as ordinary kriging procedure applied to the data on air emission submitted by installation operators to the NED database can provide the estimation of error of the emission assessment. Results of analysis carried out for fuel combustion installations located in Silesian Province (2011), presented in Figures 12.6 and 12.7 and kriging standard error estimations (presented in Figures 12.6b and 12.7b).

Analysis carried out with using original input data on emission of air pollutants (SO_x and NO_x)

Derived from the NED database is one of the first this type analysis conducted in Poland, however the geostatistical methods are widely applied, e.g. [3, 7, 8, 9]. This type of analysis (analyzed data located in only one province) can be compared to analysis provided by Leopold [7], however unmodified procedure of analysis are not possible to apply in case of whole country due to scaling problem and other downscaling issues. Elaborated analysis (especially part on kriging estimation) highlighted several issues connected with consistency of data collected in the NED database.

REFERENCES

- 1 EMEP/EEA, 2013, Air pollutant emission inventory Guidebook, Chapter 7: Spatial emissions mapping, <http://www.eea.europa.eu/publications/emep-eea-guidebook-2013>. [28.03.2014].
- 2 European Parliament, Directive 2010/75/EU of the European Parliament and of the Council of 24 November 2010 on industrial emissions (integrated pollution

- prevention and control).
- 3 Guevara M., Baldasano JM, Aréval G., Gassó S., Martínez F., Soret A.: Development of a spatial and temporal disaggregation methodology applied to the EMEP emission inventory, in Proceedings: ACCENT-IGAC GBA June 11-13, Toulouse, France 2012.
 - 4 Journal of Laws of the Republic of Poland, Dz.U. 2009 Nr 130 poz. 1070, Law act of 17 July 2010 on greenhouse gases and air pollutants emissions management system (in Polish).
 - 5 Journal of Laws of the Republic of Poland, Dz.U. 2011 Nr 3 poz. 4, Regulation of the Minister of Environment of the Republic of Poland of 28 December 2010 on default form of the report and way of its insertion to the National Emission Database on Greenhouse Gases and Other Pollutants (in Polish).
 - 6 de Kluzenaar Y., Aherne J., Farrell E.P.: Modeling the spatial distribution of SO₂ and NO_x emissions in Ireland, *Environmental Pollution* 112, 2001, pp.171-182.
 - 7 Leopold U., Heuvelink G.B.M., Drouet L., Zachary D.S.: Modeling Spatial Uncertainties associated with Emission Disaggregation in an integrated Energy Air Quality Assessment Model, in Proceedings: International Congress on Environmental Modeling and Software Managing Resources of a Limited Planet, 6th Biennial Meeting, Leipzig, Germany 2012.
 - 8 Loriato A.G., Albuquerque T.T., Marinho R.S., Santiago A.M., Salvador N., Nascimento E.G.S., Reis Jr. N.C.: Emissions Inventory for the Metropolitan Area of Vitoria, ES, Brazil using SMOKE System, in Proceedings: Community Modeling & Analysis Center, October 28, 2013.
 - 9 Maes J., Vliegen J., Van de Vel K., Janssen S., Deutsch F., De Ridder K., Mensink C.: Spatial surrogates for the disaggregation of CORINAIR emission inventories, *Atmospheric Environment* 43, 2009, pp. 1246-1254.
 - 10 National Centre for Emissions Management, 2014, Poland's Informative Inventory Report 2014, Submission under UN ECE Convention on Long-range Transboundary Air Pollution (report approved by the Ministry of Environment)
 - 11 Ribeiro P.J.: Package geoR, <http://cran.r-project.org/web/packages/geoR/geoR.pdf>, [29.03.2014].
 - 12 RICARDO-AEA, UK Emission Mapping Methodology 2009. A report of the 2013, Ricardo-AEA/R/3105.
 - 13 Shu Y., Lam N.S.N.: Spatial disaggregation of carbon dioxide emissions from road traffic based on multiple linear regression model, *Atmospheric Environment* 45, 2011, pp. 634-640.
 - 14 Skiba U., Sozanska M., Metcalfe S., Fowler D.: Spatially Disaggregated Inventories of Soil NO and N₂O Emissions for Great Britain, *Water, Air, and Soil Pollution: Focus* 1, 2001, pp. 109–118.
 - 15 Stephenson J., Gallagher K., Holmes C.C.: Beyond kriging - Dealing with discontinuous spatial data fields using adaptive prior information and Bayesian

- Partition Modelling 2003.
- 16 Tsilingiridis G., Sidiropoulos C., Pentaliotis A., Evripidou C., Papastavros C., Mesimeris T., Papastavrou M.: A spatially allocated emission inventory for Cyprus, *Global NEST Journal* 12, 2010, pp. 99-107.
 - 17 United Nations Economic Commission for Europe, 1979, Convention on Long-Range Transboundary Air Pollution.
 - 18 Schneider P.: Downscaling emission data from a global emission inventory database for use in air dispersion model, in *Proceedings: BAQ Conference 2012 - December 2012*.
 - 19 Volfová A., Šmejkal M.: Geostatistical Methods in R, in *Proceedings: Geoinformatics*, <http://geoinformatics.fsv.cvut.cz/latex/gifce-demo.pdf>, [29.03.2014].
 - 20 Webster R., Olivier M.A.: *Geostatistics for environmental scientists* 2nd Edition, Wiley 2007.
 - 21 Zasina D., Zawadzki J.: Statistical analysis of data set on national reporting of emission of air pollutants. Part I: investigation of outliers, *Environmental Protection and Natural Resources* No. 24, Issue 3, 2013.
 - 22 Zasina D., Zawadzki J.: On application of basic geostatistical methods in air pollutant emission inventory. Case study: emission of NOX and SOX from Silesian combustion plants, in *Proceedings: 1st Conference of Doctoral Candidates in Environmental Engineering, Faculty of Environmental Engineering, Warsaw University of Technology, 2013* (in Polish).
 - 23 Zawadzki J.: *Metody geostatystyczne dla kierunków przyrodniczych i technicznych*, OWPW, Warszawa 2011.
 - 24 Zawadzki J., Zasina D.: *Emission Inventory for Lignite Based Public Power and Energy Sector – Merging Information from Various Sources, Systems Supporting Production Engineering 2013*,

DISAGGREGATION PROBLEMS USING DATA DERIVED FROM POLISH AIR POLLUTANT EMISSION MANAGEMENT SYSTEM

Abstract: This paper presents analysis carried out with using data provided by the Polish system of air pollutant emissions management – the National Emission Database (NED; data from 2011), established by the National Centre for Emissions Management at the Institute of Environmental Protection – National Research Institute.

The NED database was established primarily for purposes of country-level air pollutant emission management also financial matters in environmental management (environmental taxes). The first analysis: [21] also the second [22] have shown that the data collected in the NED database can be treated as an effective, supporting device in national air pollutant emission inventory for purposes of fulfilling international obligations (i.e. UN ECE LRTAP Convention [17]).

In this paper are shown selected problems with emission downscaling and disaggregation for purposes of reporting with using data on emission of nitric oxides (NO_x) and sulphuric oxides (SO_x) from stationary combustion sources derived from the NED database.

Key words: National Emission Database, emission inventory, downscaling, disaggregation, Poland

PROBLEMY DYSAGREGACJI Z WYKORZYSTANIEM DANYCH POCHODZĄCYCH Z POLSKIEGO SYSTEMU ZARZĄDZANIA EMISJAMI ZANIECZYSZCZEŃ POWIETRZA

Streszczenie: W niniejszej pracy przedstawiono wyniki analizy danych o rocznej emisji (dane za rok 2011) tlenków azotu (NO_x) oraz siarki (SO_x) zgłoszonych przez operatorów stacjonarnych instalacji spalania paliw, zlokalizowanych w województwie śląskim, w ramach sprawozdawczości krajowej prowadzonej przez Instytut Ochrony Środowiska – Państwowy Instytut Badawczy, w Krajowym Ośrodku Bilansowania i Zarządzania Emisjami.

Przedstawiona analiza potwierdza użyteczność krajowego systemu zarządzania emisjami w aspekcie prowadzenia analiz związanych z dysagregacją oszacowań krajowej emisji zanieczyszczeń do powietrza.

Słowa kluczowe: Krajowa baza, inwentaryzacja emisji, dysagregacja danych, Polska

mgr inż. Damian ZASINA
Institute of Environmental Protection – National Research Institute
Warsaw 00-805, 132/134 Chmielna Str.
e-mail: Damian.Zasina@kobize.pl

dr hab. inż. Jarosław ZAWADZKI, prof. PW
Warsaw University of Technology
Faculty of Environmental Engineering
Warsaw 00-653, 20 Nowowiejska Str.
e-mail: J.J.Zawadzki@gmail.com