GUIDELINES FOR PERFORMING ATMOSPHERIC DEPOSITION ASSESSMENTS, INCLUDING A PROPOSAL FOR AN APPROPRIATE DEPOSITION MODELING TECHNIQUE TOGETHER WITH TRAINING OF GIOŚ EMPLOYEES

Task 4.
Guidelines for performing atmospheric deposition assessments, including a proposal for an appropriate deposition modeling technique

Report commissioned by the Chief Inspectorate of Environmental Protection under contract GIOŚ/ZP/129/2023/DMŚ/MFEOG dated May 24, 2023 for the project **"Strengthening of atmospheric deposition assessment in Poland based on Norwegian experience"** under the program "Environment, Energy and Climate Change", area "Climate Change Mitigation and Adaptation", financed by the European Economic Area Financial Mechanism 2014–2021

Manager
of the Air Quality Monitoring Department

mgr Julita Biszczuk-Jakubowska

The study was performed at the Institute of Meteorology and Water Management – National Research Institute by the team composed of: Julita Biszczuk-Jakubowska, Magdalena Bogucka, Agnieszka Kolanek, Mikołaj Kowal, Ewa Krajny, Ewa Liana, Katarzyna Lewandowska, Andrzej Mazur, Leszek Ośródka, Michał Pobudejski

WARSAW, DECEMBER 1, 2023

Contents

1. Introduction

Guidelines for performing atmospheric deposition assessments constitute another important part of the process of optimizing atmospheric deposition measurements in Poland started in 2021.

2. Aim of the project

The aim of the project is to develop guidelines for performing atmospheric deposition assessments, including a proposal for an appropriate deposition modeling technique together with training of Chief Inspectorate of Environmental Protection (GIOŚ) employees, using Norwegian experience as part of the implementation of the project titled "Strengthening of atmospheric deposition assessment in Poland based on Norwegian experience" under the program "Environment, Energy and Climate Change", area "Climate Change Mitigation and Adaptation", financed by the European Economic Area Financial Mechanism 2014–2021.

The development of guidelines for performing atmospheric deposition assessments using selected methods will enable to obtain the spatial distribution of selected deposition indicators across Poland. An appropriate modeling technique aims to expand knowledge of the spatial distribution of deposition indicators especially in regions (areas) where no measurements are available. The guidelines are intended to make it possible to prepare an annual report on the atmospheric deposition assessment and to track multi-year trends and changes. Implementation of the project should contribute to the design and establishment of a coherent national system for performing deposition assessments, responding to current national and international needs.

3. Basis for the study

The basis for the study is contract no. GIOŚ/ZP/129/2023/DMŚ/MFEOG, concluded on May 24, 2023 between the State Treasury – Chief Inspectorate of Environmental Protection (GIOŚ) and the Institute of Meteorology and Water Management – National Research Institute (IMGW-PIB).

4. Scope of the study

The scope of this study includes recommendations for deposition modeling techniques, based on the results of Tasks 1–3, comments made by the Norwegian partner NILU and GIOŚ and discussions held during the training on November 28–29, 2023. The study includes guidelines for performing atmospheric deposition assessments, with an indication of an appropriate deposition modeling technique.

4.1. Introduction

As part of Task 1, a literature review of statistical interpolation methods for atmospheric deposition assessment was performed. Calculations were performed to obtain the spatial distribution of selected four indicators: SO_4^{2-}, NO_3^-, NH_4^+, Cd for the base year 2021. Three interpolation methods were used, i.e. two kriging methods (ordinary and universal) and a local polynomial interpolation method. For comparison, interpolation was also performed using the inverse distance weighted method, as a method used to date in

annual assessments of atmospheric deposition in Poland. Calculations were performed for 3 different grid sizes: 4 x 4 km, 10 x 10 km and 50 x 50 km. The uncertainty of selected interpolation methods was estimated.

A base year for tests related to finding the best statistical interpolation methods for wet deposition in Poland was selected taking into account the thermal and precipitation conditions and wet deposition of the past several years. On the basis of the data from 22 precipitation chemistry measurement stations from the period 2017–2022, the variability of deposition in these years was analyzed, the concentration ranges were examined and the completeness of the data was assessed. Thermal and precipitation conditions were also assessed in consecutive years and against the multi-year period.

As a result of the analysis, 2021 was selected as the base year used for testing spatial distributions.

As part of the implementation of Task 1, the ordinary kriging (OK) method was selected as the appropriate statistical method for atmospheric deposition assessments. For homogeneous data, such as the tested annual average deposition values of SO_4^{2-}, NO_3^- and NH_4^+, it is recommended to use ordinary kriging without transformation and with the spherical variogram function. In the case of data involving outliers, as in the case of the tested average annual cadmium deposition, it is recommended to use ordinary kriging with logarithmic transformation and the spherical variogram function. In order to improve the accuracy of the interpolation results obtained, it is recommended to additionally use data on the precipitation level using the cokriging method.

A comparison of the spatial distributions of atmospheric deposition obtained with the currently used inverse distance weighted (IDW) method and the ordinary kriging method recommended for the future has not revealed significant differences. Higher concentrations around measurement points were noticeable when using the IDW method, especially when neighboring sites had very different concentration levels. If the number of precipitation measurement points was very large and at the same time many times larger than the number of sampling points for measuring precipitation chemistry, the results should be similar for both methods.

The calculations for selecting the optimum grid size, performed in Task 1 for 3 different grid sizes (50 x 50 km, 10 x 10 km, 4 x 4 km), were based on the properties of the input data, namely the number of measurement points and the size of the interpolated area. Both for the 22 currently operating precipitation chemistry measurement stations and for the planned 19 measurement sites that will be part of the modernized precipitation chemistry monitoring system in Poland, a 10 x 10 km grid is recommended.

As part of the implementation of Task 2, a literature review was performed of mathematical models of chemical transport which can be used to obtain the spatial distribution of deposition indicators across Poland. The EMEP model was selected as the recommended model for future use.

As part of Task 3, preliminary draft guidelines for performing atmospheric deposition assessments were developed. Following their presentation, suggestions were received from the Norwegian partner NILU and GIOŚ, which were incorporated into the project update as additional opportunities to improve the quality of atmospheric deposition

assessment in Poland. The updated draft guidelines were presented during a training course for GIOŚ employees organized on November 28–29, 2023, with the participation of the Norwegian partner NILU. The comments made by the trainees have been incorporated into the final version of the guidelines for performing atmospheric deposition assessments, presented below, including a proposal for an appropriate deposition modeling technique.

4.2. Guidelines for performing atmospheric deposition assessments by statistical interpolation method

ESRI's ArcGIS software package, release 10.2.1, along with the Geostatistical Analyst (GA) tools extension, which includes a set of tools to examine spatial data and generate surface models using statistical and deterministic methods, can be used for spatial analysis of atmospheric deposition.

Before performing statistical interpolation, the data should be analyzed for homogeneity, i.e. the presence of outliers. If the data have a positively skewed distribution (Figure 4.1) and there are several very large values, a transformation should be performed to reduce variability and normalize the data (ESRI, 2021).

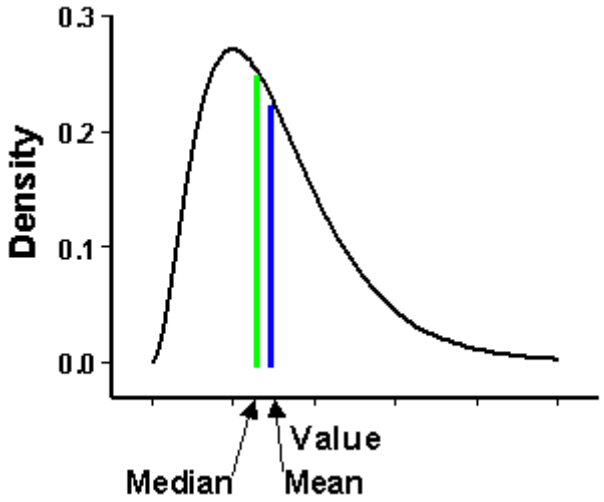


Figure 4.1 Example of input data distribution for which it is recommended to use logarithmic transformation (source: ESRI, 2021)

For homogeneous data it is recommended to use ordinary kriging without transformation and with the spherical variogram function. In the case of data involving outliers, it is recommended to use ordinary kriging with logarithmic transformation and the spherical variogram function.

If any doubt arises about the use of the above-mentioned transformation, it is recommended to perform two interpolations (with and without transformation) and choose the better one, using the appropriate statistical parameters (mean prediction error – ME, root mean square error – RMSE, average standard error – ASE, mean standardized prediction error – MSE, root mean square standardized prediction error – RMSSE), and to assess the prediction match with the measurement data.

Maps of spatial distributions of atmospheric deposition should be made for a grid size of 10 x 10 km.

The quality of the interpolation performed is significantly affected by the number of measurement points, which determines the magnitude of the error and the uncertainty of measurements. It is recommended to use as many available measurement points as possible. Due to the correlation of atmospheric deposition with precipitation level, in order to improve the accuracy of the results obtained, it is recommended to additionally use data on the precipitation level, e.g. from the IMGW-PIB station network, and in order to obtain their spatial distribution – to use the cokriging method. The precipitation data is available free of charge at: https://danepubliczne.imgw.pl/.

Atmospheric deposition should be assessed based on monthly and annual average maps.

It is recommended to standardize the scale by adopting fixed ranges in equal steps, constant for successive research years. This will make it possible to compare deposition values by month and by year. The recommended ranges used to generate spatial distribution maps for monthly and annual average values are detailed in the following tables (4.1 – 4.2).

Table 4.1 Recommended scale ranges of spatial distribution maps for annual average values of selected components

SO_4^{2-} [kg/ha SO_4^{2-}]	NO_3^{2-} [kg/ha N]	NH_4^+ [kg/ha N]	Cd [g/ha Cd]	Precipitation [mm/m^2]
0.00 - 2.60	0.00 - 0.80	0.00 - 1.30	0.000 - 0.870	0.0 - 157.0
2.61 - 5.20	0.81 - 1.70	1.31 - 2.60	0.871 - 1.740	157.1 - 314.0
5.21 - 7.80	1.71 - 2.60	2.61 - 3.90	1.741 - 2.610	314.1 - 471.0
7.81 - 10.40	2.61 - 3.50	3.91 - 5.20	2.611 - 3.480	471.1 - 628.0
10.41 - 13.00	3.51 - 4.40	5.21 - 6.50	3.481 - 4.350	628.1 - 785.0
13.01 - 15.60	4.41 - 5.30	6.51 - 7.80	4.351 - 5.220	785.1 - 942.0
15.61 - 18.20	5.31 - 6.20	7.81 - 9.10	5.221 - 6.090	942.1 - 1,099.0
18.21 - 20.80	6.21 - 7.10	9.11 - 10.40	6.091 - 6.960	1,099.1 - 1,256.0
20.81 - 23.40	7.11 - 8.00	10.41 - 11.70	6.961 - 7.830	1,256.1 - 1,413.0
> 23.41	> 8.01	> 11.71	> 7.831	> 1,413.1

Table 4.2 Recommended scale ranges of spatial distribution maps for monthly average values of selected components

SO_4^{2-} [kg/ha SO_4^{2-}]	NO_3^{2-} [kg/ha N]	NH_4^+ [kg/ha N]	Cd [g/ha Cd]
0.00 - 0.80	0.00 - 0.20	0.00 - 0.40	0.000 - 0.310
0.81 - 1.60	0.21 - 0.40	0.41 - 0.80	0.311 - 0.620
1.61 - 2.40	0.41 - 0.60	0.81 - 1.20	0.621 - 0.930
2.41 - 3.20	0.61 - 0.80	1.21 - 1.60	0.931 - 1.240
3.21 - 4.00	0.81 - 1.00	1.61 - 2.00	1.241 - 1.550
4.01 - 4.80	1.01 - 1.20	2.01 - 2.40	1.551 - 1.860
4.81 - 5.60	1.21 - 1.40	2.41 - 2.80	1.861 - 2.170
5.61 - 6.40	1.41 - 1.60	2.81 - 3.20	2.171 - 2.480
6.41 - 7.20	1.61 - 1.80	3.21 - 3.60	2.481 - 2.790
> 7.21	> 1.81	> 3.61	> 2.791

The following Figures (4.2 – 4.21) present sample maps of the spatial distributions of atmospheric deposition (annual and monthly average values) of the four indicators tested. Maps of annual average values (except for cadmium) were made using the OK method without data transformation, using the spherical variogram function. Maps of annual average values for cadmium were developed using the logarithmic transformation. Maps of monthly average values were made with or without logarithmic transformation, depending on the presence of outliers in a given month.

The maps presented (Figures 4.2 – 4.21) indicate that the use of fixed scale ranges can in some cases result in a lack of differentiation in the values of interpolation results for low variability data (Figures 4.3, 4.4, 4.8, 4.9, 4.13, 4.14, 4.18, 4.19). When the values in the input dataset are dispersed, it is possible to obtain information about the variability of deposition values in the analyzed area (Figures 4.2, 4.5, 4.6, 4.7, 4.10, 4.11, 4.12, 4.15, 4.16, 4.17, 4.20, 4.21).

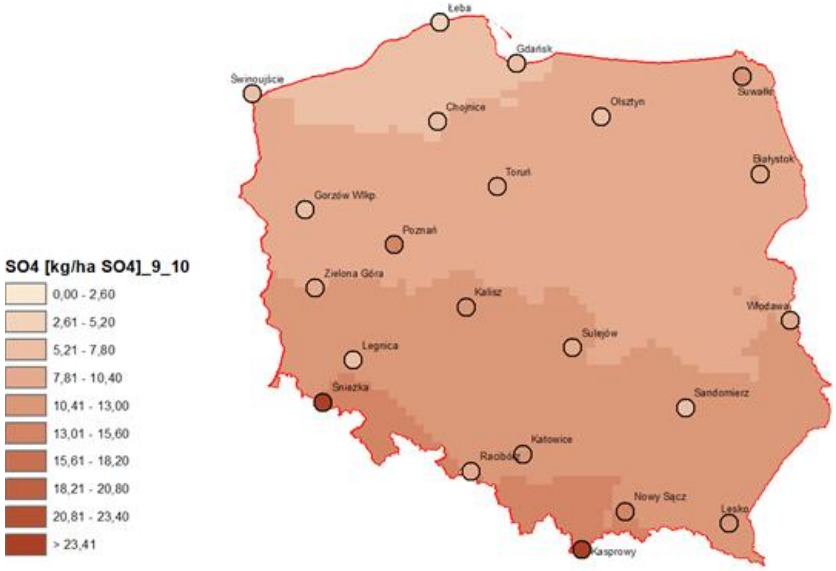


Figure 4.2 Spatial distribution of atmospheric deposition of sulphate loads [kg/ha SO_4^{2-}] in Poland in 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

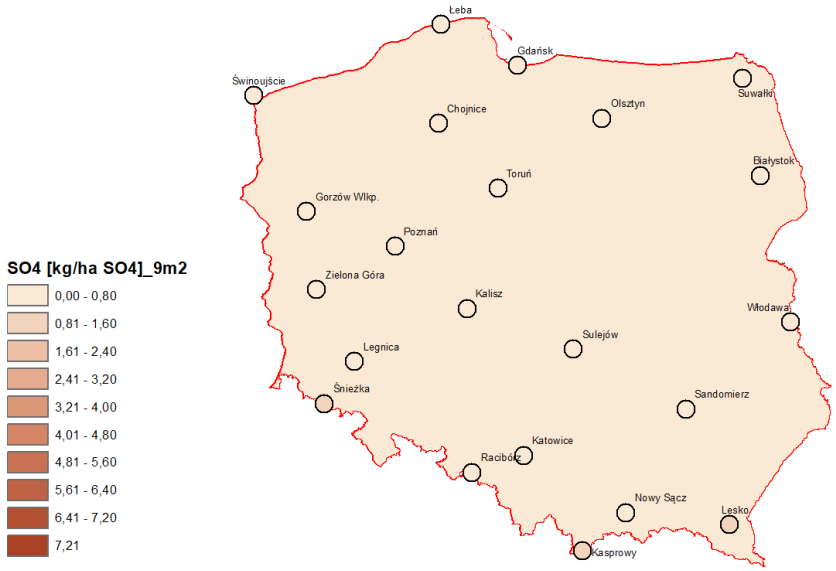


Figure 4.3 Spatial distribution of atmospheric deposition of sulphate loads [kg/ha SO_4^{2-}] in Poland in February 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

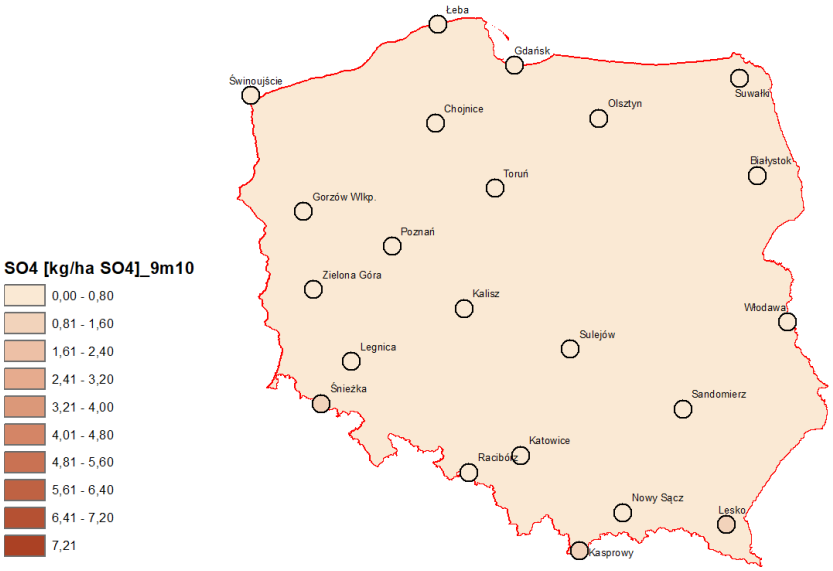


Figure 4.4 Spatial distribution of atmospheric deposition of sulphate loads [kg/ha SO_4^{2-}] in Poland in October 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

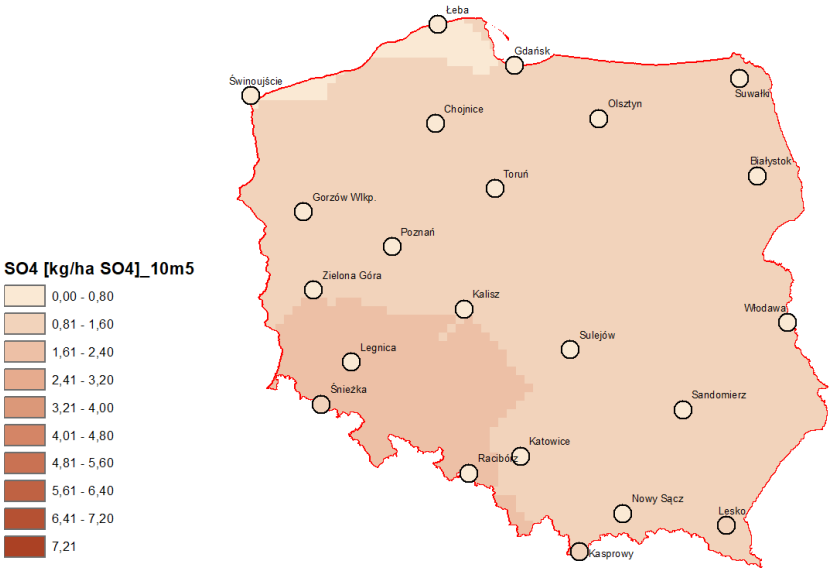


Figure 4.5 Spatial distribution of atmospheric deposition of sulphate loads [kg/ha SO_4^{2-}] in Poland in May 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

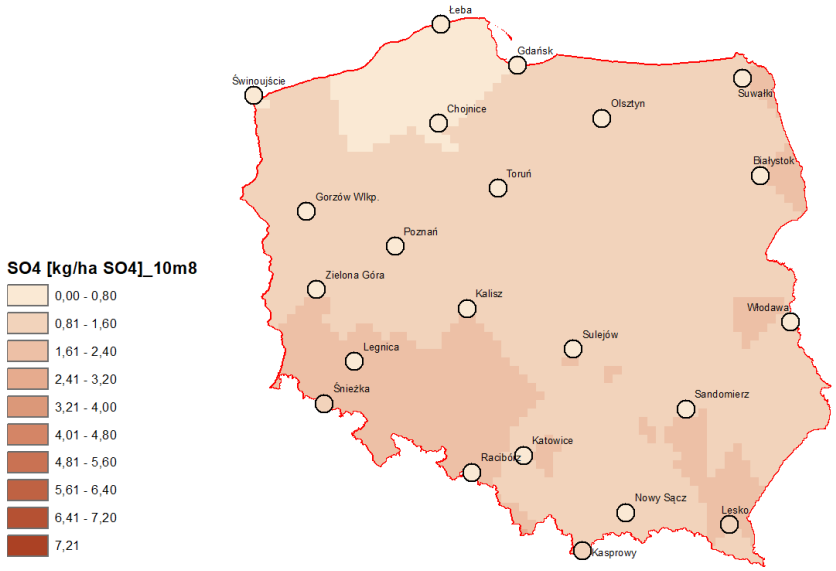


Figure 4.6 Spatial distribution of atmospheric deposition of sulphate loads [kg/ha SO_4^{2-}] in Poland in August 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

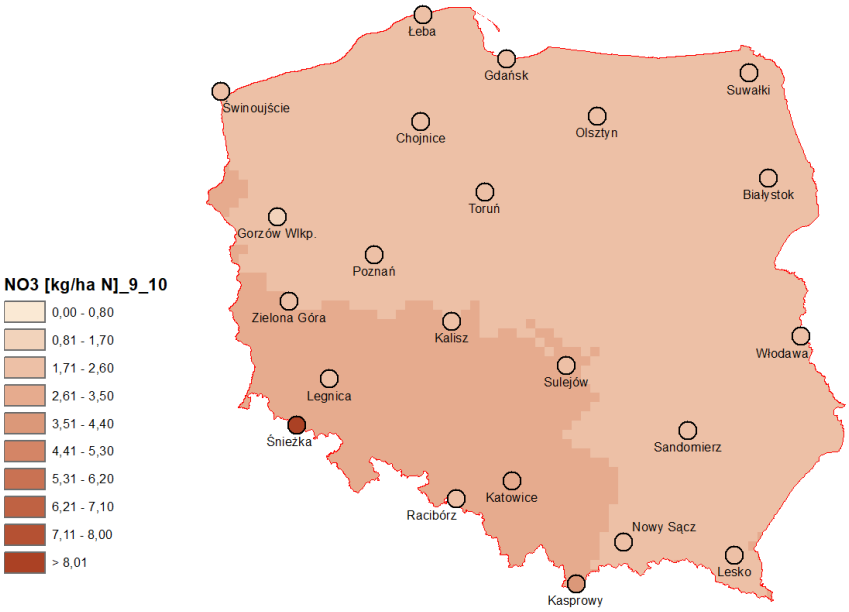


Figure 4.7 Spatial distribution of atmospheric deposition of nitrate loads [kg/ha N] in Poland in 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

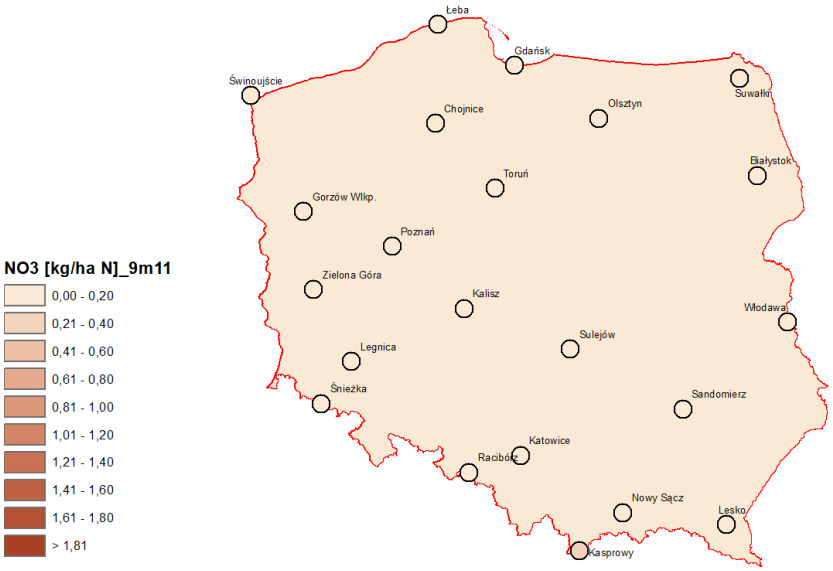


Figure 4.8 Spatial distribution of atmospheric deposition of nitrate loads [kg/ha N] in Poland in November 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

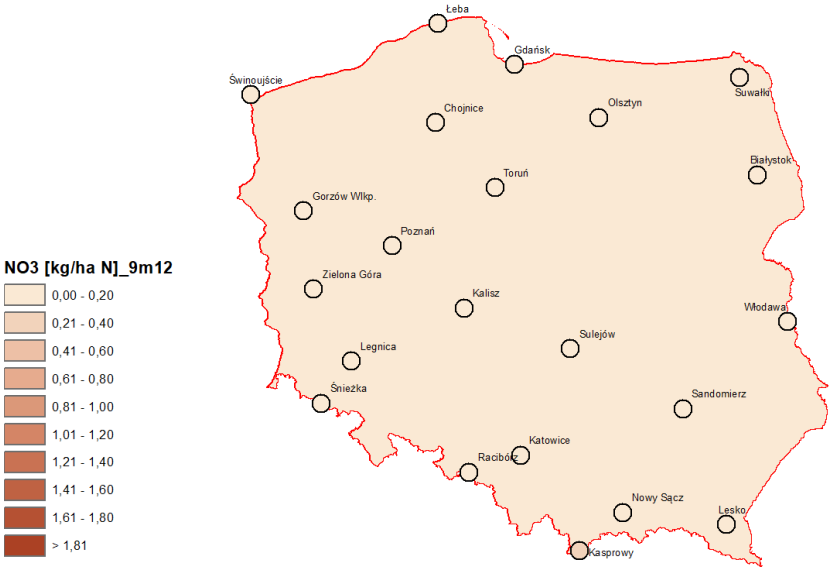


Figure 4.9 Spatial distribution of atmospheric deposition of nitrate loads [kg/ha N] in Poland in December 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

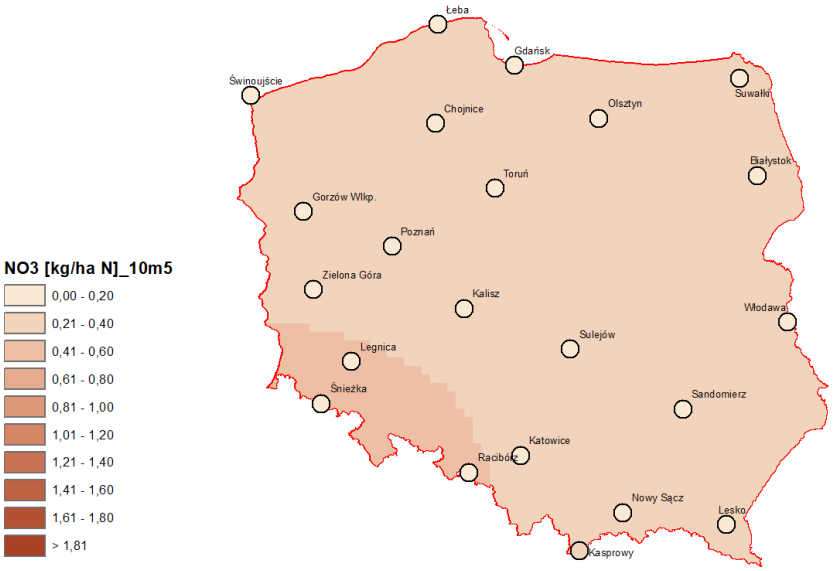


Figure 4.10 Spatial distribution of atmospheric deposition of nitrate loads [kg/ha N] in Poland in May 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

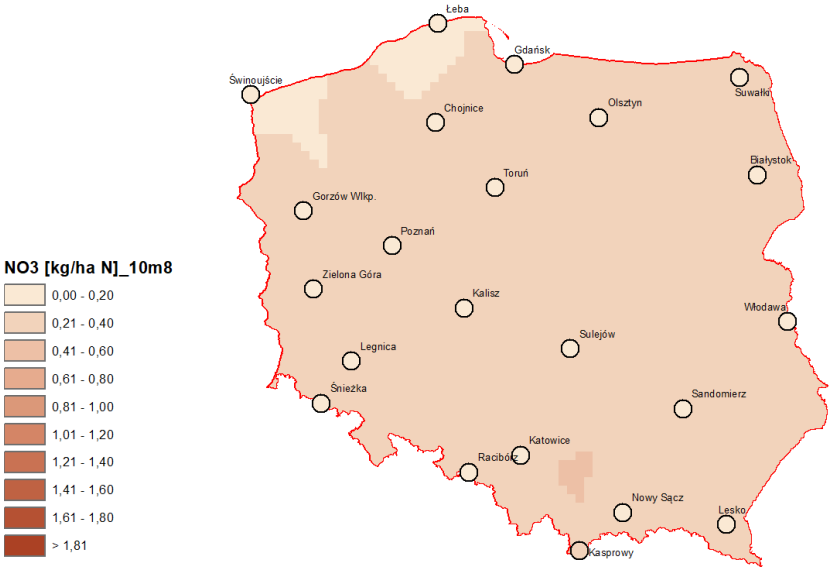


Figure 4.11 Spatial distribution of atmospheric deposition of nitrate loads [kg/ha N] in Poland in August 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

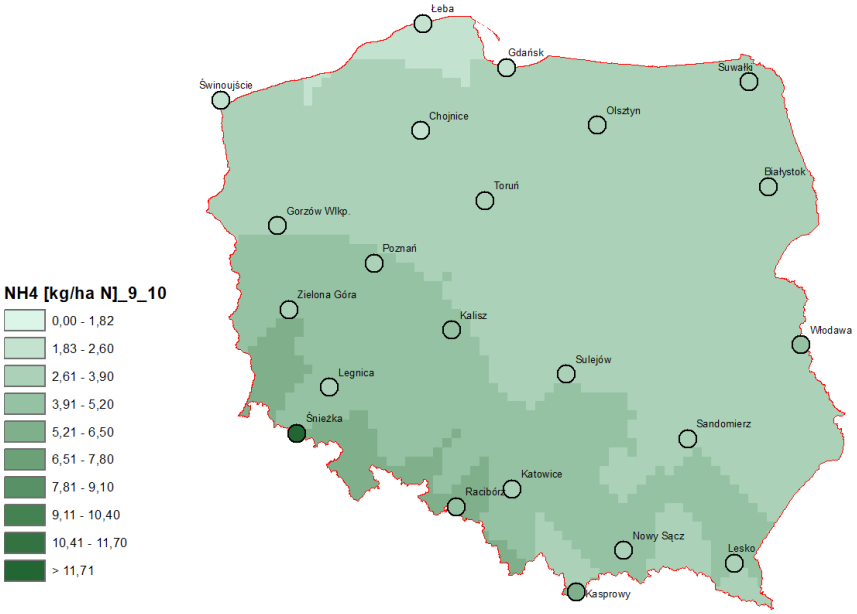


Figure 4.12 Spatial distribution of atmospheric deposition of ammonium nitrogen loads [kg/ha N] in Poland in 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

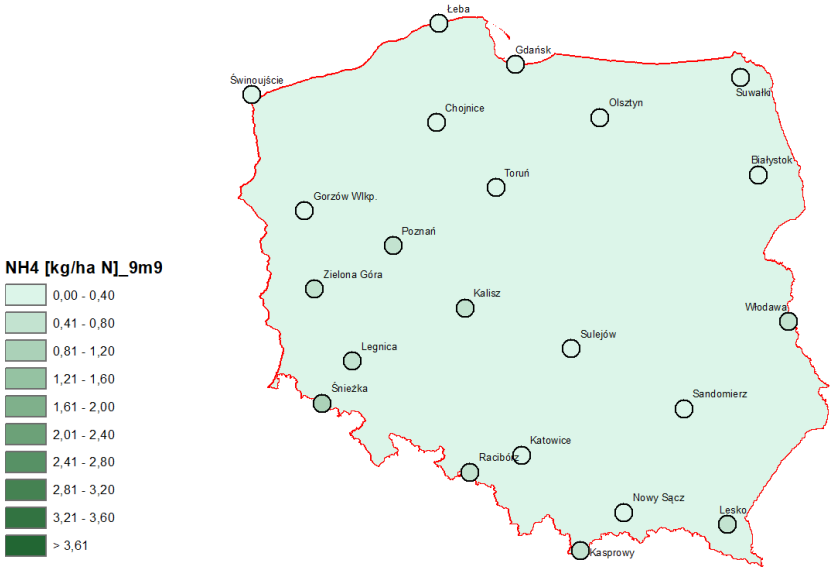


Figure 4.13 Spatial distribution of atmospheric deposition of ammonium nitrogen loads [kg/ha N] in Poland in September 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

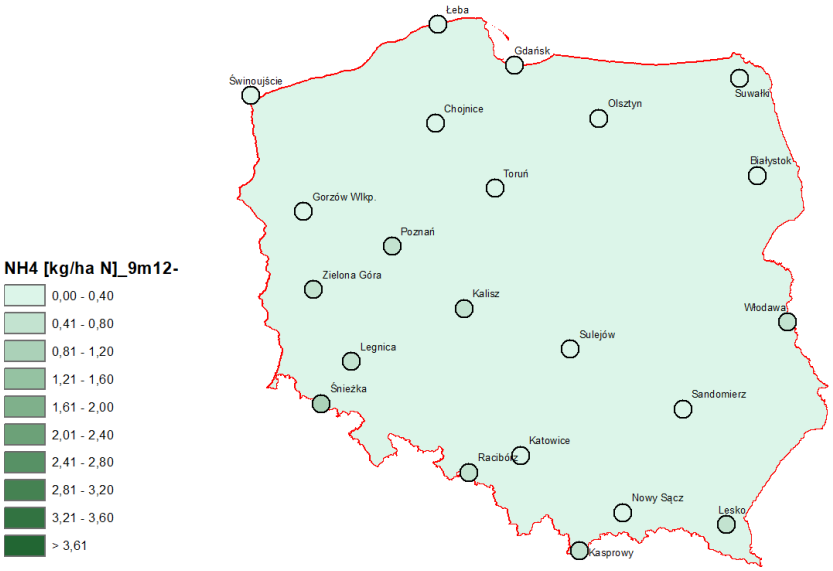


Figure 4.14 Spatial distribution of atmospheric deposition of ammonium nitrogen loads [kg/ha N] in Poland in December 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

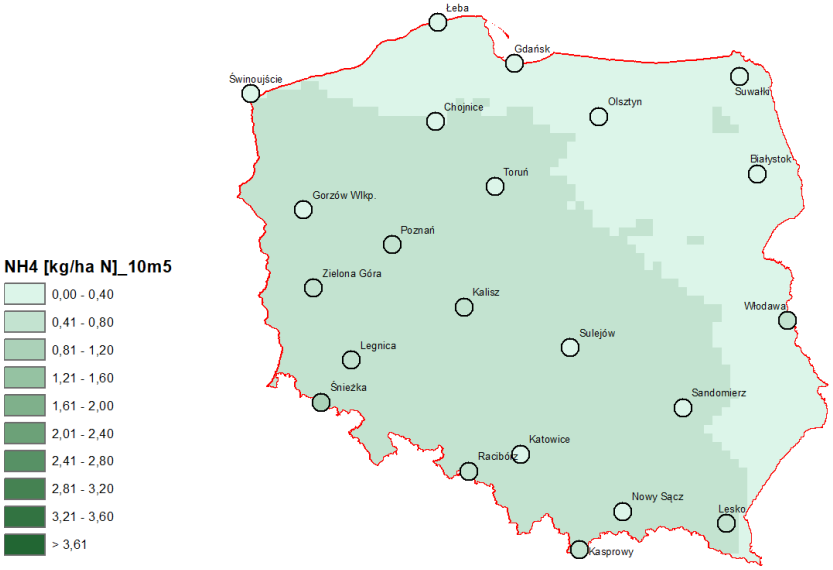


Figure 4.15 Spatial distribution of atmospheric deposition of ammonium nitrogen loads [kg/ha N] in Poland in May 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

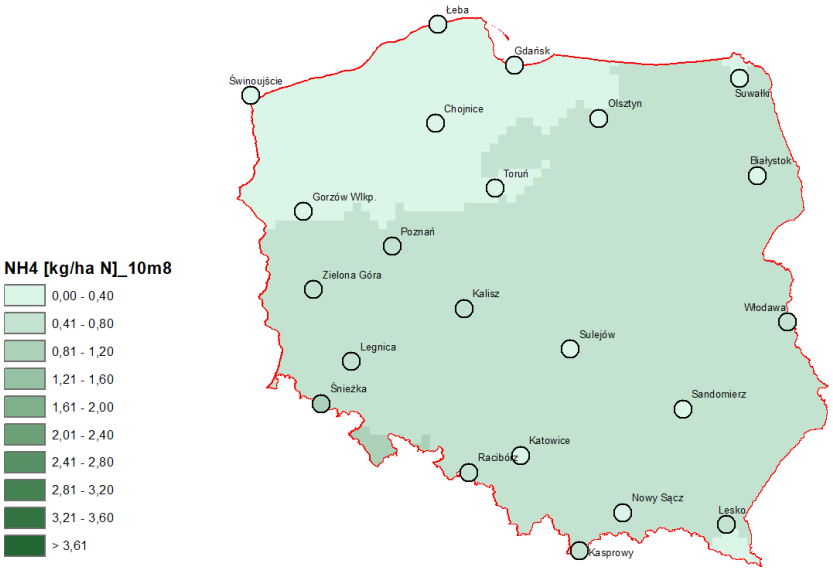


Figure 4.16 Spatial distribution of atmospheric deposition of ammonium nitrogen loads [kg/ha N] in Poland in August 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

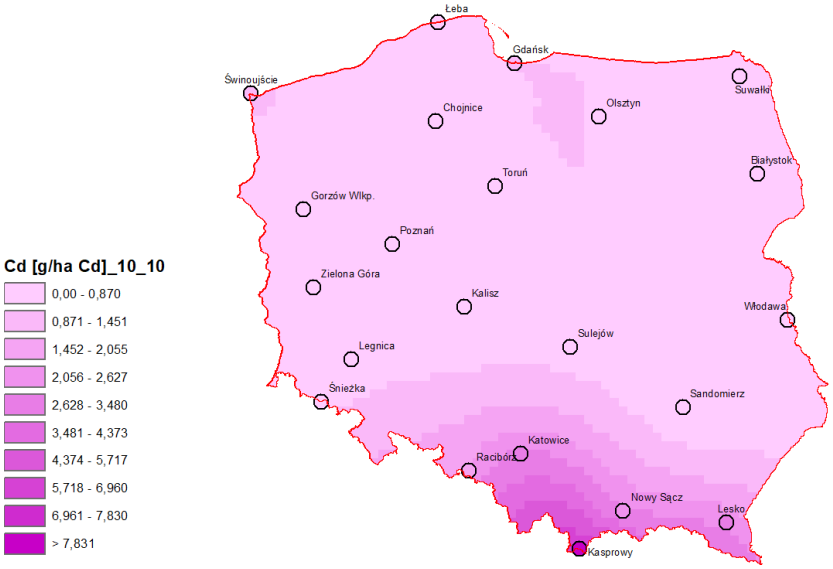


Figure 4.17 Spatial distribution of atmospheric deposition of cadmium loads [g/ha Cd] in Poland in 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

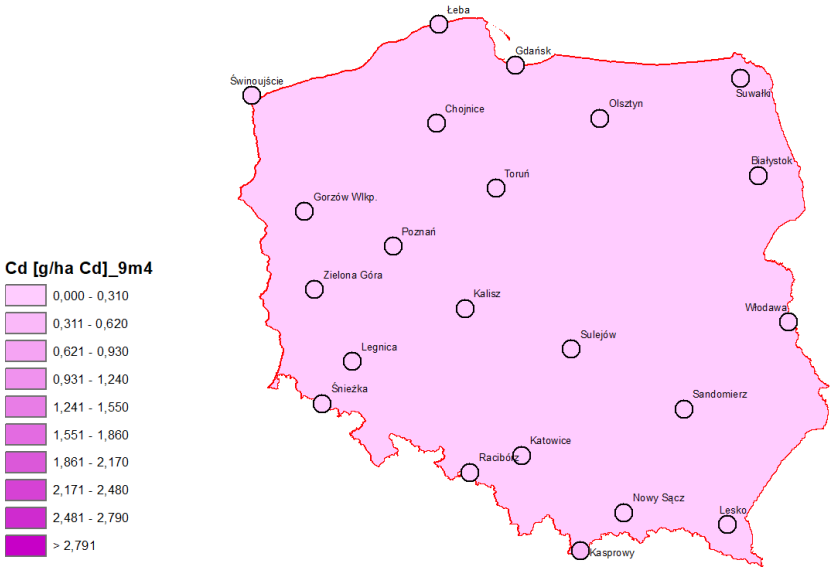


Figure 4.18 Spatial distribution of atmospheric deposition of cadmium loads [g/ha Cd] in Poland in April 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

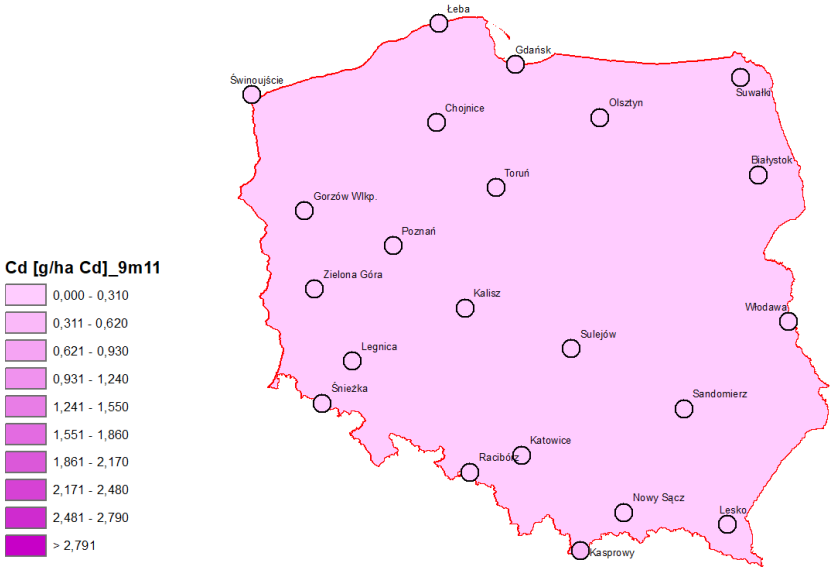


Figure 4.19 Spatial distribution of atmospheric deposition of cadmium loads [g/ha Cd] in Poland in November 2021 obtained using the ordinary kriging method without transformation and with the spherical variogram function for a grid size of 10 x 10 km

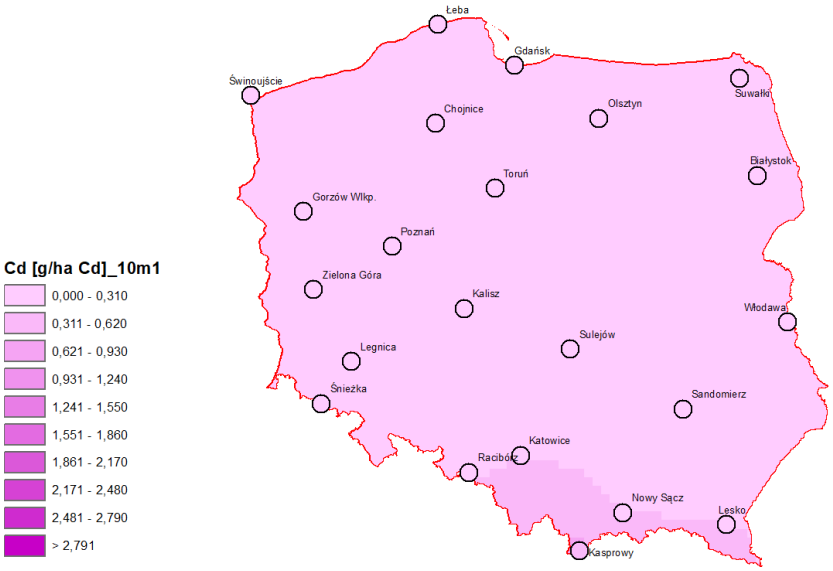


Figure 4.20 Spatial distribution of atmospheric deposition of cadmium loads [g/ha Cd] in Poland in January 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

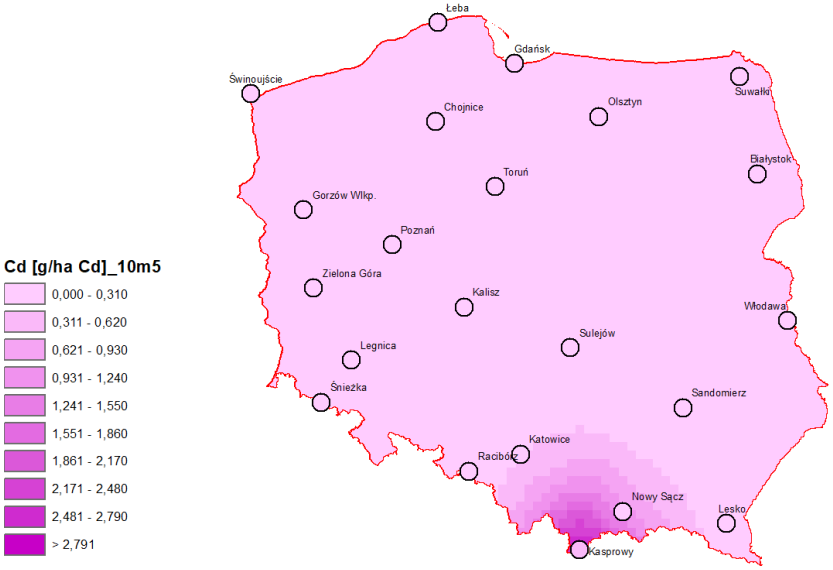


Figure 4.21 Spatial distribution of atmospheric deposition of cadmium loads [g/ha Cd] in Poland in May 2021 obtained using the ordinary kriging method with logarithmic transformation and with the spherical variogram function for a grid size of 10 x 10 km

The proposed methodology for assessing atmospheric deposition has many advantages, but also some limitations. The advantages of using the kriging method over the IDW method include:

- determination of the most probable value in a given layer (raster cell), and assessment of the expected interpolation error,
- calculation of statistical characteristics of the interpolation error allows the selection of the best parameters used in the process,
- assessment of the probability of occurrence at a given location of values larger or smaller than a preset threshold value,
- possibility of extending the analysis to include other variables in the interpolation (cokriging).

Limitations to the use of the kriging method are:

- the small number of measurement points and their irregular distribution,
- no representativeness of the data,
- the presence of outliers can have a major effect on interpolation results for the areas in which they were measured,
- large spatial variation in the data can cause high interpolation errors.

The development of a fixed color scale with specific values for a selected number of ranges also brings advantages and limitations when analyzing spatial distributions. In terms of advantages, it should be noted that a fixed color scale improves the readability of the maps and allows visual comparison of average annual distribution maps on a multi-year basis. On the other hand, this approach does not allow to differentiate extreme values. The scale of annual variability does not seem suitable for assessing monthly variability due to the different range of values present. In order to correctly visualize monthly distributions, color scales should be established for each month separately. In summary, a variable scale allows visualization of extreme values, while a fixed scale facilitates comparison of the same spatial-temporal distributions.

4.3. Guidelines for performing atmospheric deposition assessments using a mathematical chemical transport model

For the implementation of Task 2, in accordance with the Terms of Reference, an analysis of literature material on mathematical models of chemical transport was performed, which made it possible to develop an appropriate recommendation in this regard.

An important criterion for selecting a chemical transport model for assessing atmospheric deposition in Poland should be its widespread use in Europe. This ensures that the chemical modeling input is reliable and updated annually, as well as that the results obtained in Poland and other European countries are comparable. Therefore, the natural choice would be to use the EMEP model.

The open-source versions of the EMEP model facilitate insight into model assumptions, parameterizations, input data requirements and model code, encourage dialogue and collaboration with the modeling community, enable custom model runs, and provide insight into how to run different scenarios. The EMEP model meets all the modeling quality criteria

specified in the HTAPII, AQMEII3 and EURODELTATrends (EDT) comparison experiments. The annual update of the emission input to this model, prepared on the basis of emission data provided regularly by most European countries, is extremely important. It is essential that the data provided is prepared in these countries based on a uniform methodology, which ensures its reliability and comparability. Each year, based on the emission data thus prepared and provided from each country, the EMEP CEIP (EMEP Center on Emission Inventories and Projections) prepares an up-to-date emission database for air quality modeling at a resolution of 0.1° x 0.1° adapted to the EMEP grid. This data, along with documentation on the method of its preparation and a description of the grid, is available for download at https://www.ceip.at/the-emep-grid/gridded-emissions. Emission data in the EMEP CEIP database is broken down into 11 SNAP sectors. Alternatively, emissions are also available for GNFR sectors or the 19-sector GNFR_CAMS system. The model can also be run using a 13-sector system with any other grid emission files, as long as the emissions are assigned to SNAP or GNFR sectors. As a standard, the model is fed with meteorological data from the Integrated Forecasting System (IFS) model of the European Centre for Medium-Range Weather Forecasts (ECMWF) with a resolution of 0.1° x 0.1°. In recent years, the EMEP model has increasingly been run with WRF meteorology. In the model, the map projection and resolution of the model output are automatically adjusted to the meteorological data input files. Flexibility in terms of map projection and resolution, model domain and modeling interval has been provided to make the model easier to use. It is also convenient that the input files in which the preferred parameter and performance options are specified are in easily editable ASCII format.

A limitation of the EMEP MSC-W (European Monitoring and Evaluation Program Meteorological Synthesizing Centre – West) model is that it is highly demanding in terms of input (meteorological) data volume and computational resources.

The involvement of many research teams in the work related to the EMEP system guarantees that new modeling methods will be developed and implemented in the system.

4.4. Analysis of additional solutions proposed by the Norwegian partner to improve the quality of atmospheric deposition assessments in Poland

In the case of interpolation by regression kriging using data from the currently operating 22 precipitation chemistry measurement stations located in Poland and data from the EMEP model as supporting data in this process, the following steps were taken to assess the feasibility of such interpolation.

Acquisition of model data

EMEP model data was acquired using the dataset indicated below:

https://thredds.met.no/thredds/catalog/data/EMEP/2023_Reporting/catalog.html?dataset=EMEP/2023_Reporting/EMEP01_rv5.0_year.2021met_2021emis.nc

Data from the following variables was used for the tests:

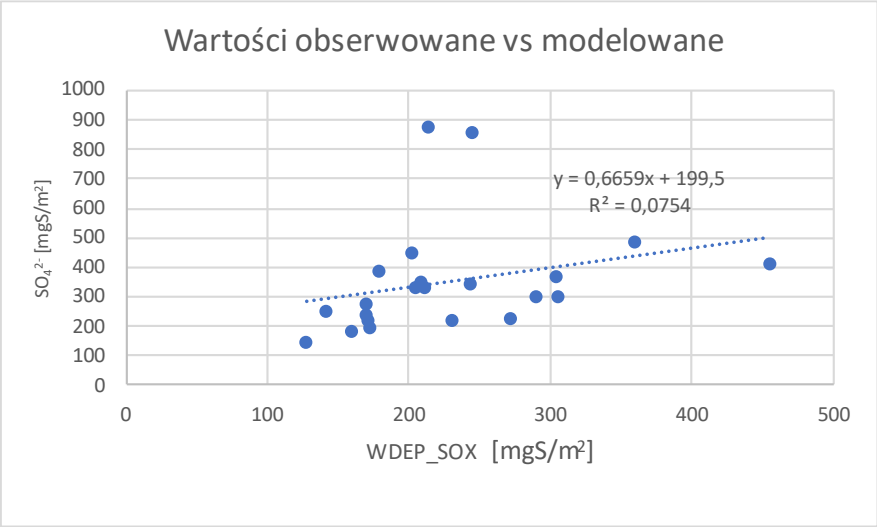
- WDEP_SOX – wet deposition of oxidized sulfur
- WDEP_OXN – wet deposition of oxidized nitrogen
- WDEP_OXN – wet deposition of reduced nitrogen

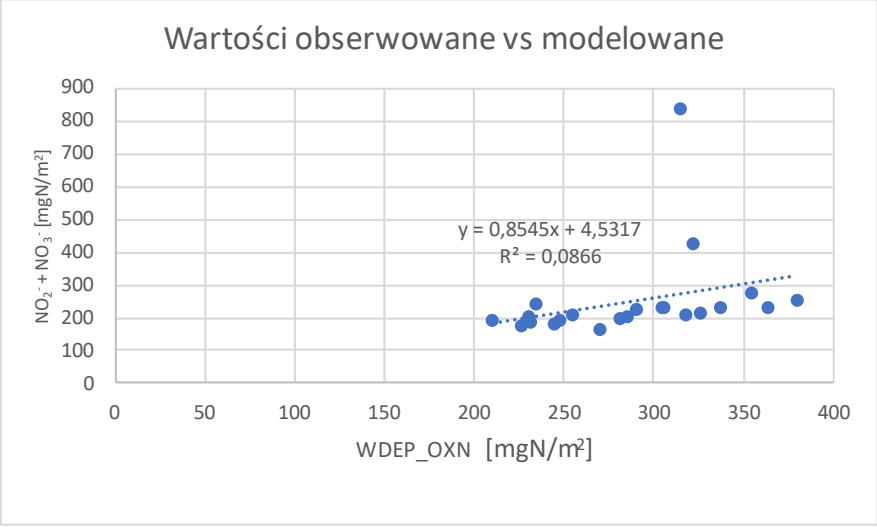
Data preparation and processing

The data was converted to the same deposition units (e.g. mgS/m^2) and prepared for use as input data for GIS software.

Checking the correlation between observation and model data

The correlation between model and observation variables was checked. The following charts (Figure 4.22) show the correlation of the observed and modeled values.





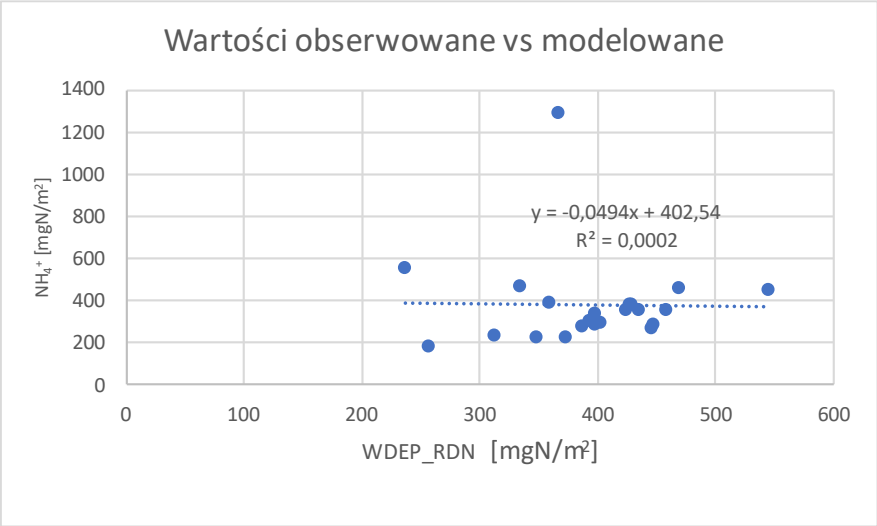
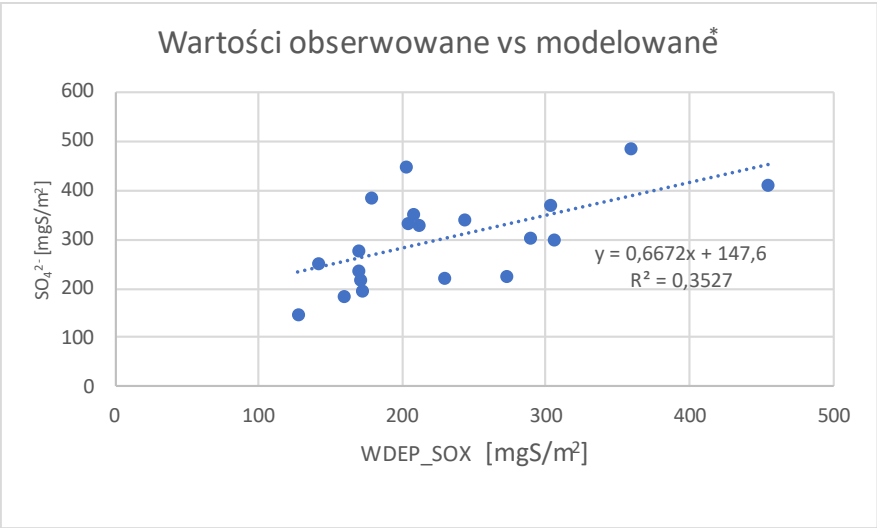
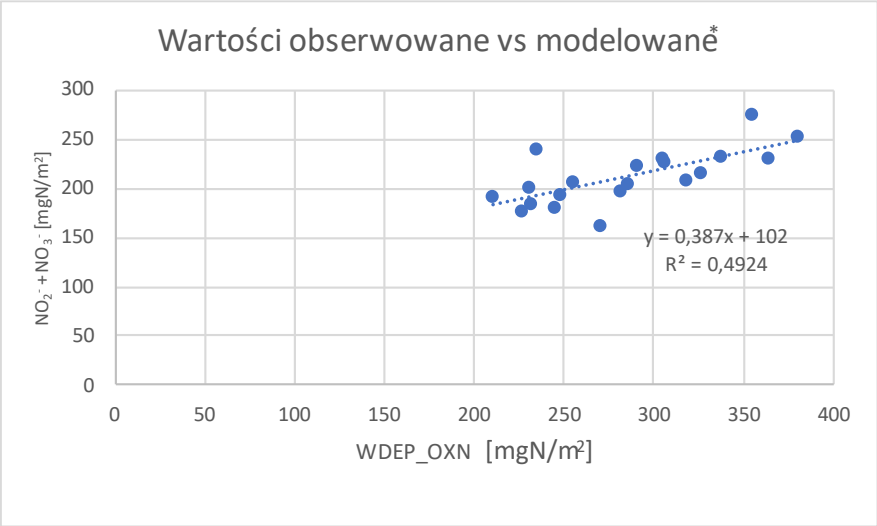


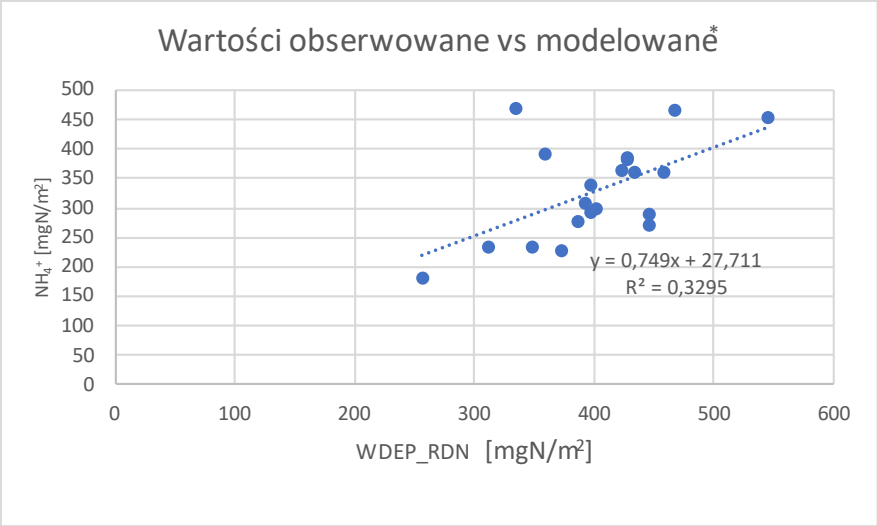
Figure 4.22 Correlation chart of observed wet deposition values and corresponding EMEP model data for all components tested

In all the above charts (Figure 4.22), the two points (in the case of WDEP_RDN it is one point) which are markedly different from the others represent the values for high-mountain stations. These stations will not be part of the newly designed system, the only high-mountain station will be the new station in Karkonosze (1,327 m ASL).

The potential removal of outliers significantly improves the correlation coefficient.







* excluding data for high-mountain stations

Figure 4.23 Correlation chart of observed wet deposition values and corresponding EMEP model data for all components tested, excluding data for high-mountain stations

Regression analysis using the OLS technique

The Ordinary Least Squares (OLS) technique was used to analyze the regression and determine the linear equation describing the correlation between the observation and model data using a tool available in the ArcGIS package.

The output generated with the OLS tool includes the Shapefile object class layer symbolized using OLS residuals, statistical results and diagnostics. They were used to perform regression analysis and interpret the OLS results.

Performance of the model is assessed using the R-squared and adjusted R-squared coefficients. It is recommended to use the adjusted R-squared coefficient, which reflects the complexity of the model (number of variables) and is therefore always lower than the R-squared value.

For the cases surveyed on the data from 22 stations, the low values of the coefficients indicate that there is no correlation between the values of the dependent

variable (observed values) and the independent variable (model values) for OLS. Removal of outliers resulted in an improved correlation coefficient, indicating a better correlation between the explanatory and response variable.

Table 4.3 Summary of statistical parameters for assessing performance of the models for all components tested

Component	Dependent value	Explanatory value	Number of observations	R-squared	Adjusted R-squared
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	0.0754	0.0292
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	0.0866	0.0409
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	0.0002	-0.0498
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	0.3527	0.3168
OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	0.4924	0.4642
RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	0.3295	0.2923

The explanatory variable was assessed by means of a coefficient, probability or reliable probability value. For the cases surveyed on the data from 22 stations for the SOX and OXN components, the values and sign of the coefficient are correct, while for RDN the sign associated with the coefficient is negative, so the correlation is negative, which is incorrect in this case. This is due to the presence of outliers. After their removal, the values and sign of the coefficients were correct.

Probability values measure the statistical significance of a coefficient. They indicate whether the effect of the explanatory variable on the response can be attributed to randomness. Low probability values indicate higher statistical significance. If a variable is not significant, it does not help the model.

For all the surveyed components measured at 22 stations, the probability values of the explanatory variable are quite high and mean that the coefficient is not statistically significant. The situation is different when outliers are removed, which resulted in a reduction in the probability value, meaning that the coefficient is statistically significant.

Table 4.4 Summary of statistical parameters for assessing the relevance of the explanatory variable used in the models for all components tested

Component	Dependent value	Explanatory value	Number of observations	Coefficient	Probability	Reliable probability
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	0.6659	0.2161	0.0219
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	0.8545	0.1837	0.0413
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	-0.0494	0.9460	0.9215
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	0.6672	0.0058	0.0012

OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	0.3870	0.0006	0.0002
RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	0.7490	0.0081	0.0046

The combined F-statistic was used to assess the overall statistical significance of the model. It is only reliable if the Koenker statistic (BP) is not statistically significant, which is the case for the tested cases.

For all the components surveyed on the data from 22 stations, the p-value indicates that the model is not statistically significant, while the opposite is true when using data from 20 measurement stations.

Table 4.5 Summary of statistical parameters for assessing the overall statistical significance of the models for all components tested

Component	Dependent value	Explanatory value	Number of observations	Combined F-statistic	p-value
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	1.6320	0.2161
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	1.8962	0.1837
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	0.0047	0.9460
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	9.8098	0.0058
OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	17.4592	0.0006
RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	8.8463	0.0081

The Koenker statistic (BP) was used to assess stationarity. This is a test to determine whether the explanatory variables in the model have a consistent relationship with the dependent variable in both geographical and data space. A p-value of less than 0.05 indicates statistically significant heteroskedasticity (variation in the relationship between predicted values and changes in the explanatory variable as the size of the dependent variable changes) or non-stationarity (mean and variance values of the explanatory variable are not constant across the surveyed area).

As shown in the table below, in all cases the p-value indicates that the test is not statistically significant and there is no evidence that the relationship is not consistent.

Table 4.6 Summary of statistical parameters for assessing heteroskedasticity and stationarity of the models for all components tested

Component	Dependent value	Explanatory value	Number of observations	Koenker (BP)	p-value
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	0.0014	0.9700
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	0.5268	0.4679
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	0.3632	0.5467
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	0.0606	0.8056
OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	0.1824	0.6694
RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	1.0914	0.0521

The Jarque-Bera statistic was used to assess model bias, which checks whether the residuals have a normal distribution. The p-value for the explanatory variable indicates that the test is statistically significant, meaning that the residuals do not have a normal distribution and the model's predictions are biased.

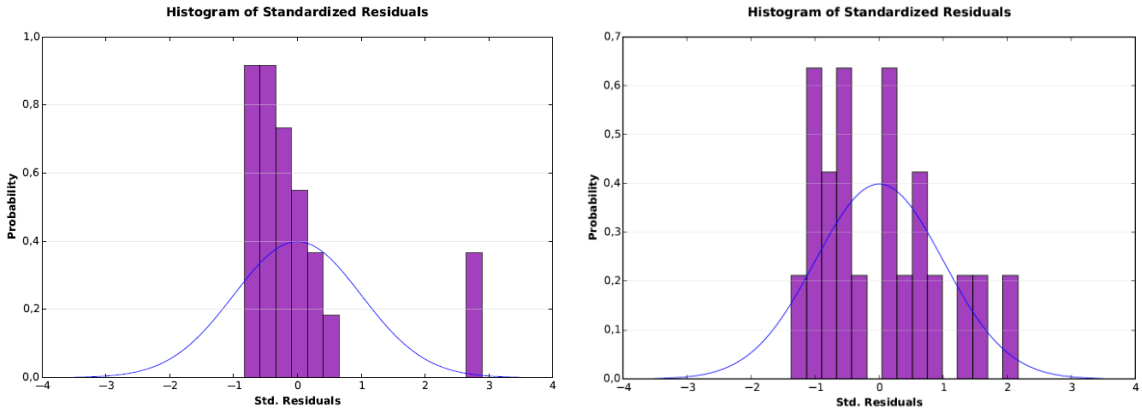
For all the components surveyed on the data from 22 stations, the distribution of the standardized residual values differs significantly from a normal distribution, indicating a bias in the model. The situation is different when data from 20 measurement stations is used, whose distribution is closer to a normal distribution.

Table 4.7 Summary of statistical parameters for assessing the similarity of the distribution of regression residuals to a normal distribution for all components tested

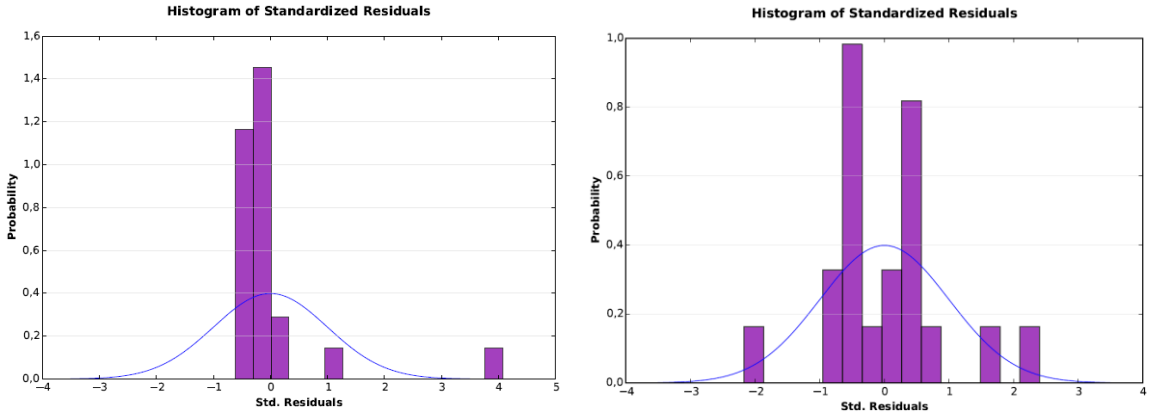
Component	Dependent value	Explanatory value	Number of observations	Jarque-Bera statistic	p-value
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	29.4671	0.0000
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	183.9517	0.0000
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	146.8600	0.0000
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	1.2911	0.5244
OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	1.8031	0.4059
RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	5.9083	0.0521

The histograms below (Figure 4.24) show how closely the distribution of standardized residuals matches the normal curve. For data from 22 stations (histograms on the left below), the appearance of the histogram deviates significantly from the normal curve, indicating that the model may be biased. The histogram based on data from 20 stations presents a much better picture (histograms on the right below).

SOX



OXN



RDN

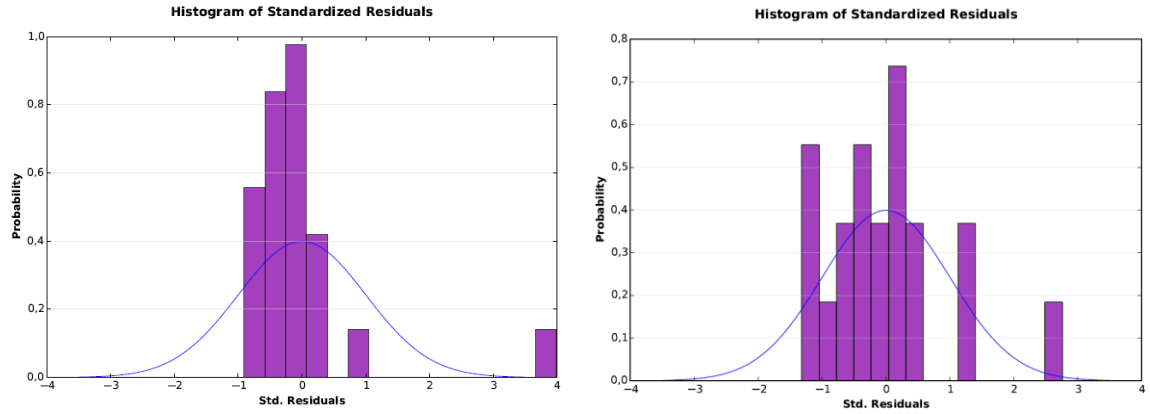


Figure 4.24 Histograms of the distribution of standardized regression residuals (for data from 22 stations – on the left, for data excluding high-mountain stations – on the right), allowing assessment of the match with the normal curve for all components tested

The spatial autocorrelation of the regression residuals was assessed to ensure their spatial randomness. A statistically significant clustering of high and low residuals indicates that the model adequately predicts underestimation and overestimation, which means that the results cannot be trusted and that the model is mis-specified.

For all cases tested, the z-score values mean that the distribution of regression residuals is not significantly different from random, as shown in the table below.

Table 4.8 Summary of statistical parameters for assessing spatial autocorrelation of regression residuals for all components tested

Component	Dependent value	Explanatory value	Number of observations	z-score	p-value
SOX	WDEP_SOX_obs	WDEP_SOX_mod	22	-0.396495	0.69174
OXN	WDEP_OXN_obs	WDEP_OXN_mod	22	-0.894986	0.370795
RDN	WDEP_RDN_obs	WDEP_RDN_mod	22	0.906611	0.364613
SOX	WDEP_SOX_obs	WDEP_SOX_mod	20	0.494404	0.621021
OXN	WDEP_OXN_obs	WDEP_OXN_mod	20	-0.104202	0.917009

RDN	WDEP_RDN_obs	WDEP_RDN_mod	20	1.444509	0.148596

Regression analysis using the OLS technique

The analysis showed that there is a poor correlation between the measured data from the current precipitation chemistry system and the data obtained from the EMEP model. The analysis also showed that for the data from 22 chemistry measurement stations and the corresponding data from the EMEP model, the explanatory variable has no effect on improving the model results, the resulting model is not statistically significant and the distribution of the residuals shows a large deviation from a normal distribution. On the other hand, the data analyzed does not show heteroskedasticity or non-stationarity, and the spatial distribution of the residuals is random.

The situation is different when two outliers corresponding to high-mountain stations are removed from the datasets. In this case, the correlation improves significantly, the coefficient of the model corresponding to the explanatory variable is statistically significant, which translates into its effect on the results obtained from the regression model, making the whole model statistically significant. The distribution of the residual values is much closer to a normal distribution, however, the data analyzed does not show heteroskedasticity or non-stationarity, and the spatial distribution of the residuals is random.

A two-step approach to calculating wet deposition distribution taking into account precipitation level data from the precipitation measurement network

Maps showing the spatial distribution of deposition of individual components were made based on data on annual average concentrations of precipitation pollutants, precipitation levels from 22 chemistry monitoring stations and precipitation levels from 162 precipitation measurement stations. The maps were made on a 10 x 10 km grid. Interpolations for SO_4^{2-}, NO_3^-, NH_4^+ were performed by ordinary kriging with the spherical variogram function. Only for the interpolation for Cd, a logarithmic transformation of the input data was used to eliminate the effect of outliers. The resulting raster layers were used to create a map showing the spatial distribution of deposition by combining an interpolation layer for the concentration of a given component with an interpolation layer for the precipitation level. The equation used to combine the raster layers is: *precipitation level x component concentration / 100 [kg/ha]*.

A comparison of maps derived from deposition data for 22 stations with maps created by combining precipitation data from 162 stations and concentration data from 22 stations reveals greater variation in deposition in the latter case. This is due to an underestimation of the effect of the precipitation volume in a specific area, which is a vital factor driving deposition volumes.

The observed effect of precipitation level was particularly evident for SO_4^{2-}, NO_3^-, NH_4^+.

For example, for SO_4^{2-}, when precipitation levels are relatively low in the Wielkopolskie Voivodeship, but concentrations are high, deposition is at an average level in relation to the rest of the country. By contrast, with relatively high precipitation levels

and low concentrations, the total pollution load is relatively high, as can be seen in the example of the Mazovia region. On maps obtained by kriging without the use of additional precipitation level data, these differences become unobservable.

The situation is similar for NH_4^+, with higher concentrations of this component in central Poland and the Włodawa area resulting in higher deposition values in these regions. This is clearly visible in the interpolation maps made by combining maps of precipitation volumes and component concentrations, but is not reflected in maps made by kriging without the use of additional precipitation level data.

In the case of NO_3^-, the analogous situation is found in the Wielkopolskie Voivodeship, where there is relatively little precipitation but the concentration of this component is higher. In the case of the complex interpolation maps, a decrease in deposition values is visible, which is not reflected in the kriging maps, where deposition values in the area are comparable to the average values for the larger area.

The same relationship can be observed using the cokriging method when deposition data from 22 stations and precipitation volume data from 162 stations is used.

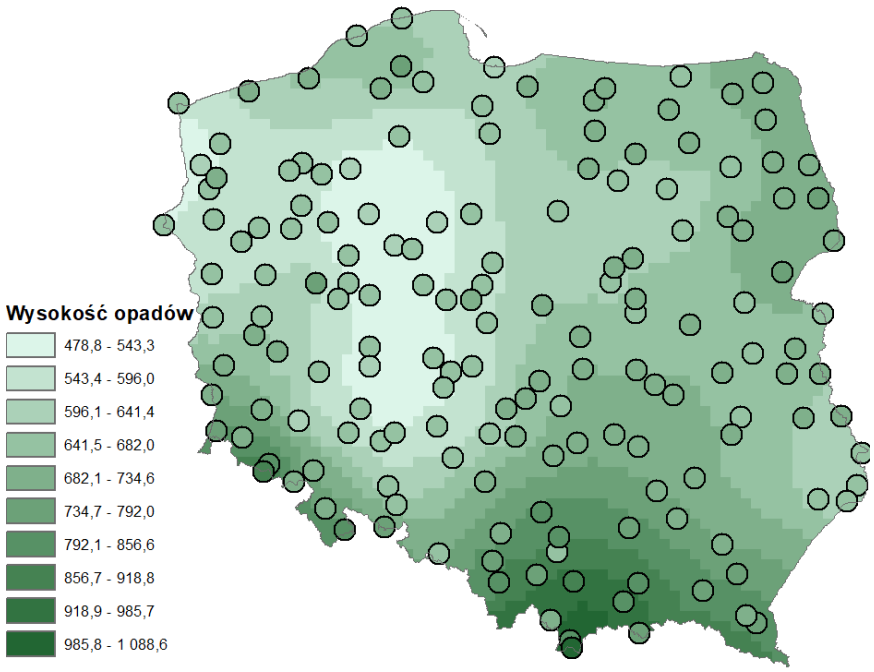


Figure 4.25 Map of precipitation level distribution for data from 162 stations, obtained by ordinary kriging with the spherical variogram function

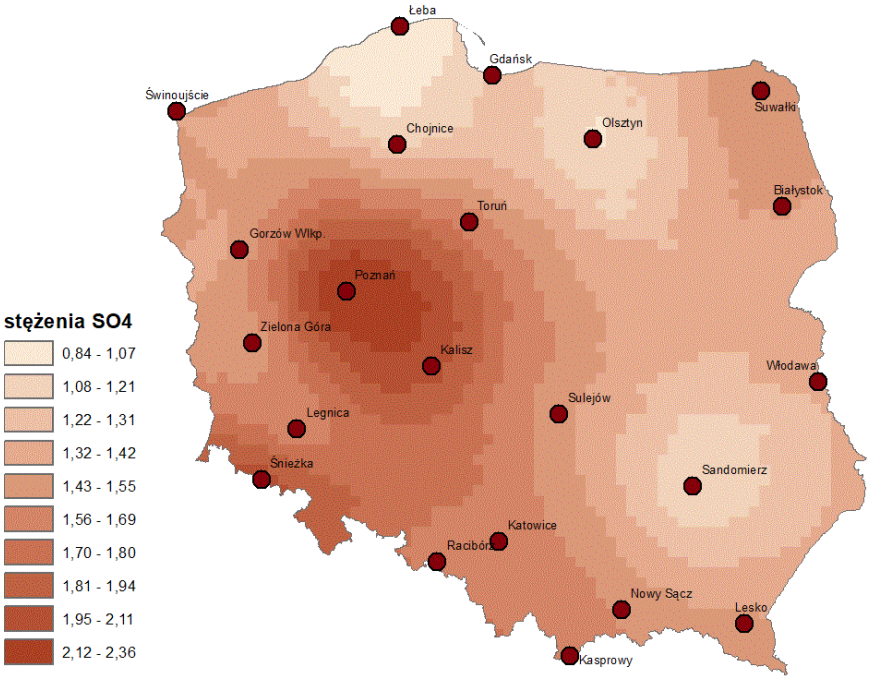


Figure 4.26 Map of SO_4^{2-} concentration distribution obtained by ordinary kriging with the spherical variogram function

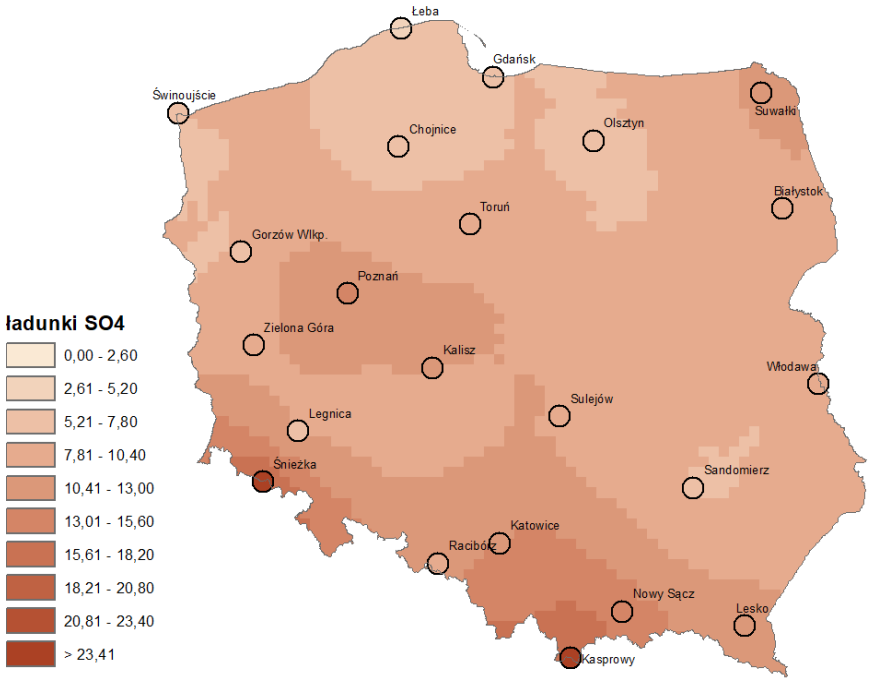


Figure 4.27 Map of SO_4^{2-} deposition distribution obtained by combining two maps: precipitation levels and SO_4^{2-} concentrations, both obtained by ordinary kriging with the spherical variogram function

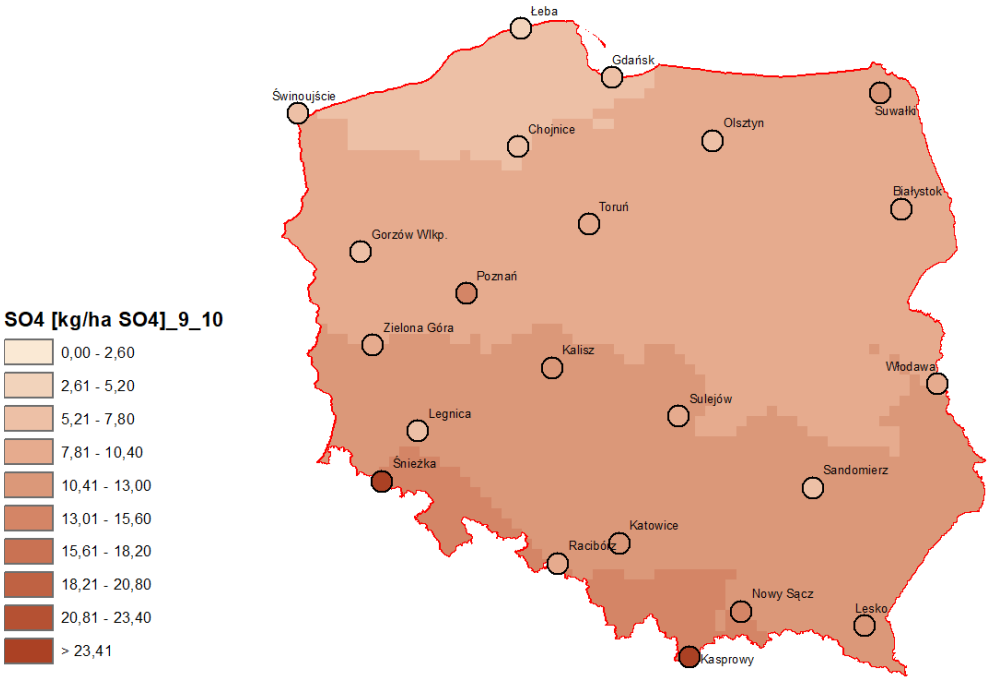


Figure 4.28 Map of SO_4^{2-} deposition distribution obtained from data from 22 stations by ordinary kriging with the spherical variogram function

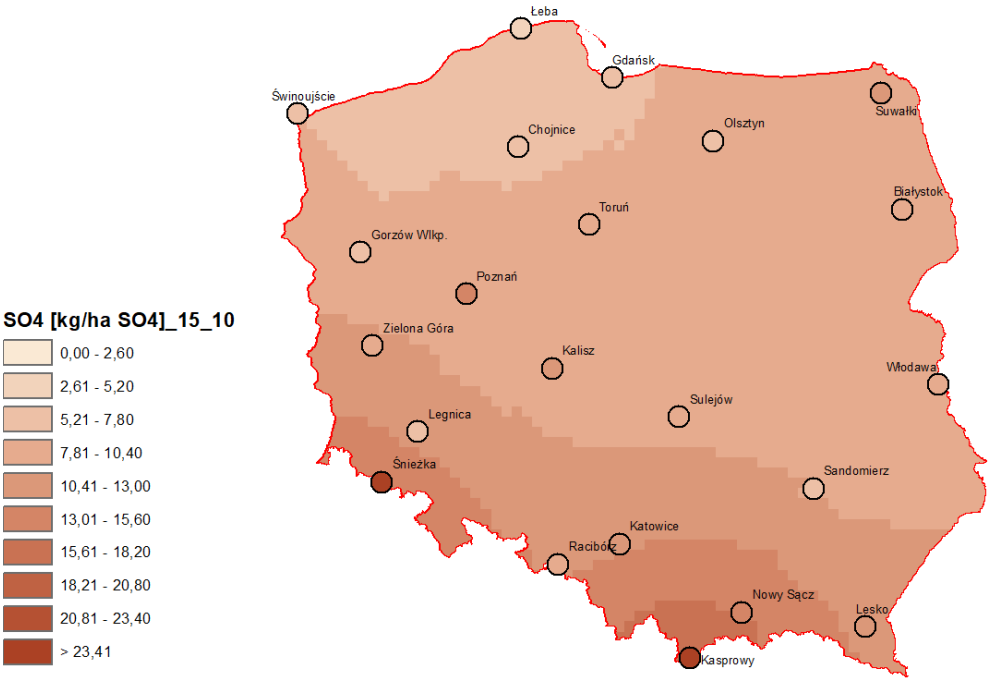


Figure 4.29 Map of SO_4^{2-} deposition distribution obtained by cokriging (with precipitation level as second variable), obtained by ordinary kriging with the spherical variogram function

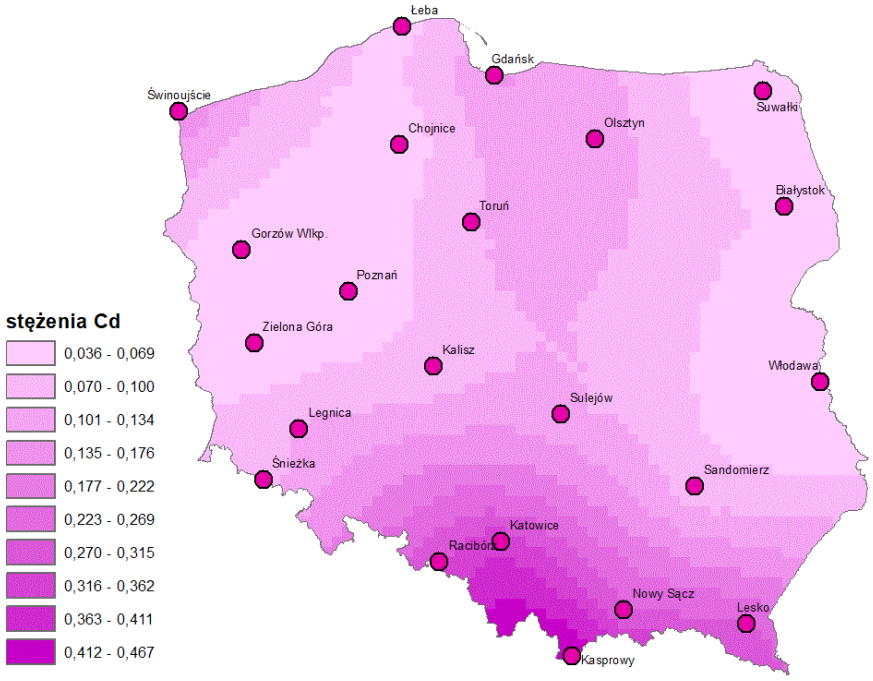


Figure 4.30 Map of Cd concentration distribution obtained by ordinary kriging with logarithmic transformation of the input data and the spherical variogram function

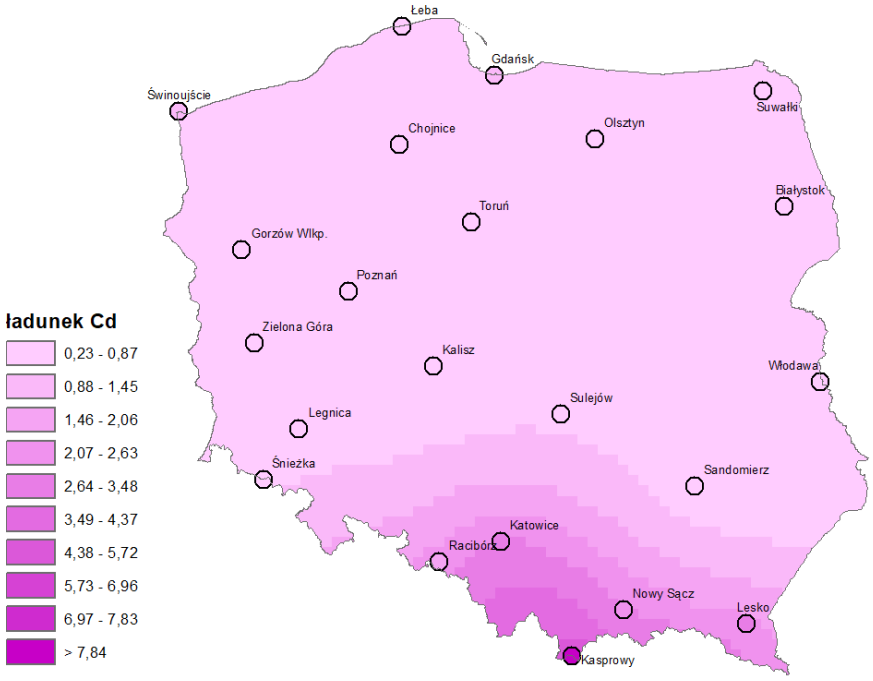


Figure 4.31 Map of Cd deposition distribution obtained by combining two maps: precipitation levels and Cd concentrations, both obtained by ordinary kriging with the spherical variogram function

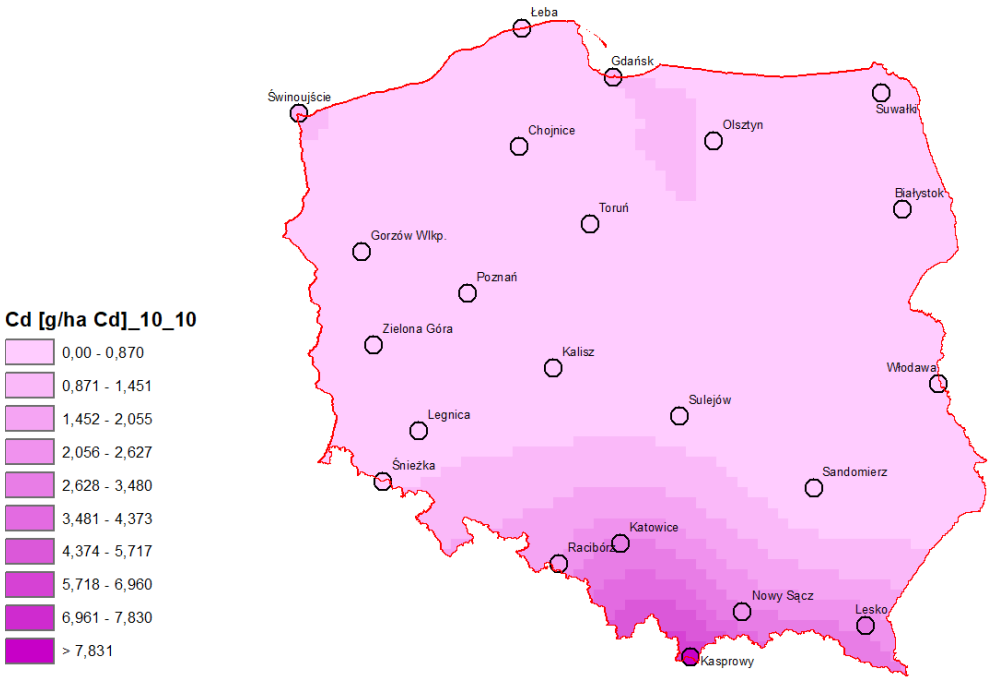


Figure 4.32 Map of Cd deposition distribution obtained by ordinary kriging with logarithmic transformation of the input data and the spherical variogram function

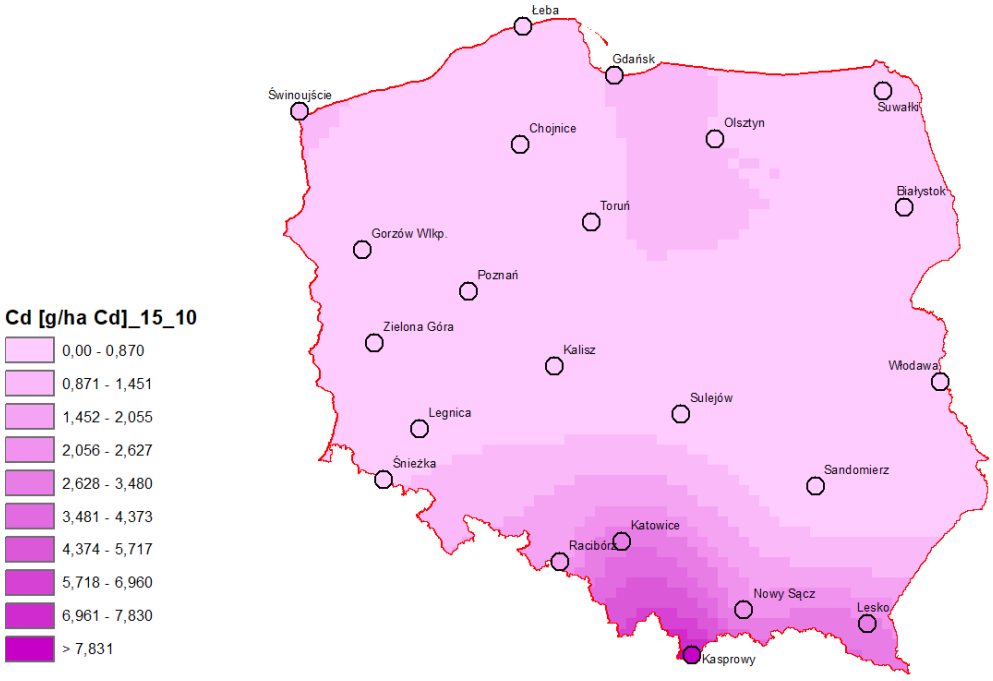


Figure 4.33 Map of Cd deposition distribution obtained by cokriging (with precipitation level as second variable), obtained by ordinary kriging
with logarithmic transformation of the input data and the spherical variogram function

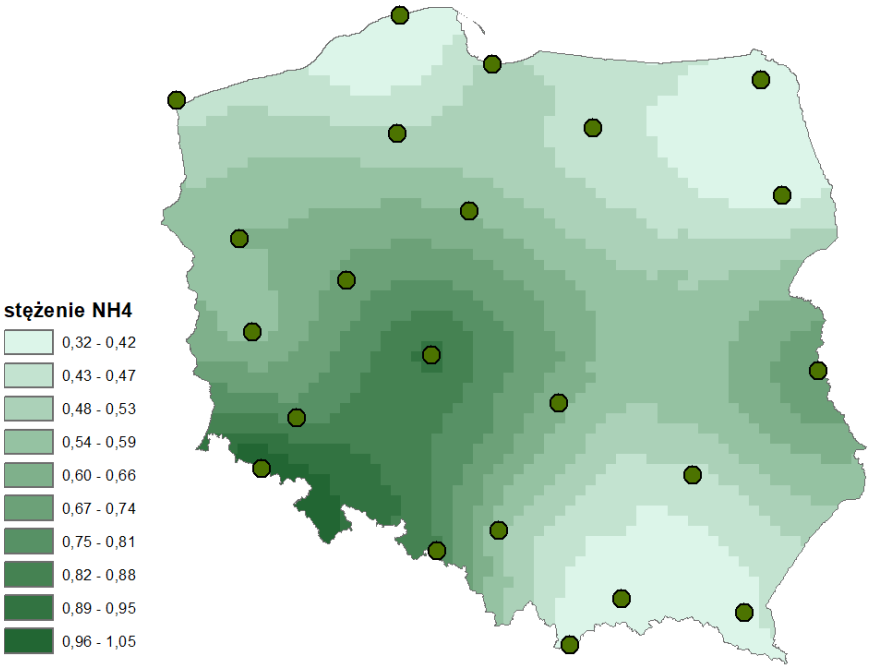


Figure 4.34 Map of NH_4^+ concentration distribution obtained by ordinary kriging with the spherical variogram function

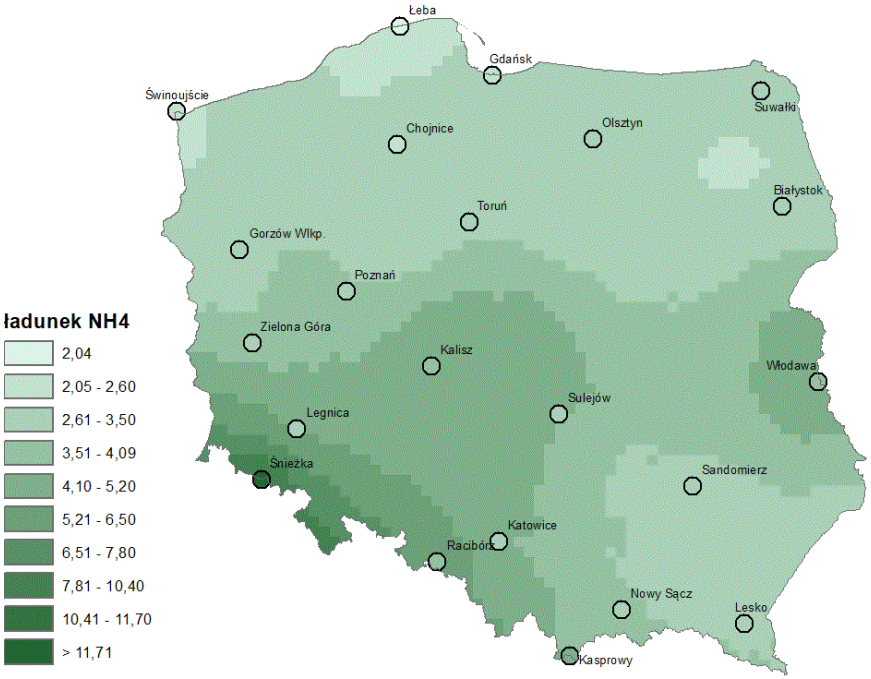


Figure 4.35 Map of NH_4^+ deposition distribution obtained by combining two maps: precipitation levels and NH_4^+ concentrations, both obtained by ordinary kriging with the spherical variogram function

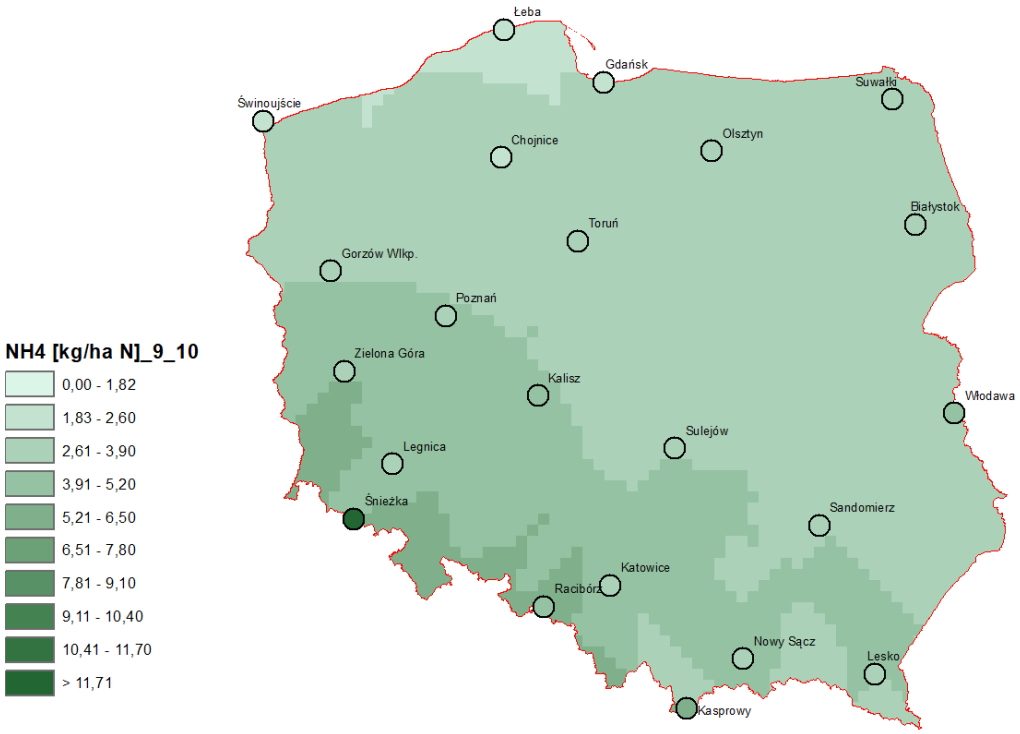


Figure 4.36 Map of NH_4^+ deposition distribution obtained from data from 22 stations by ordinary kriging with the spherical variogram function

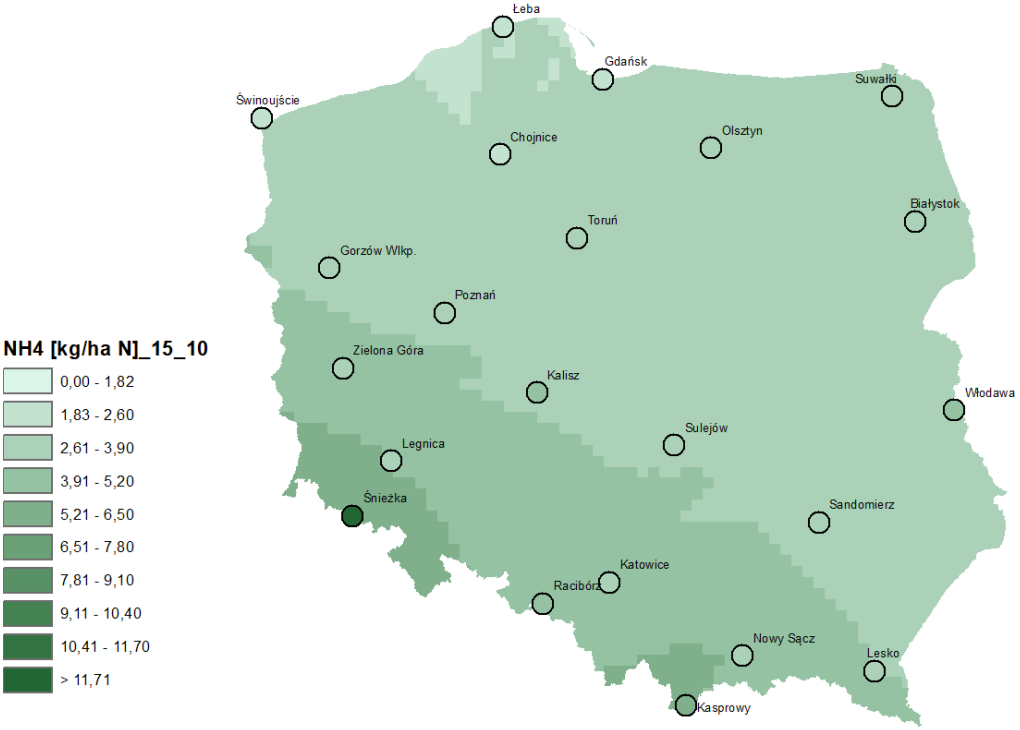


Figure 4.37 Map of NH_4^+ deposition distribution obtained by cokriging (with precipitation level as second variable), obtained by ordinary kriging with the spherical variogram function

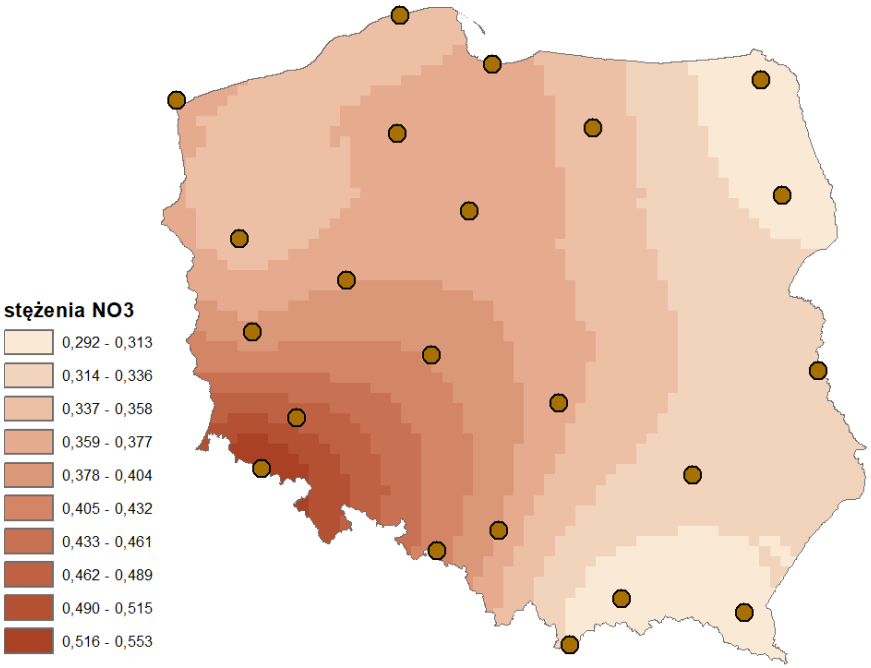


Figure 4.38 Map of NO_3^- concentration distribution obtained by ordinary kriging with the spherical variogram function

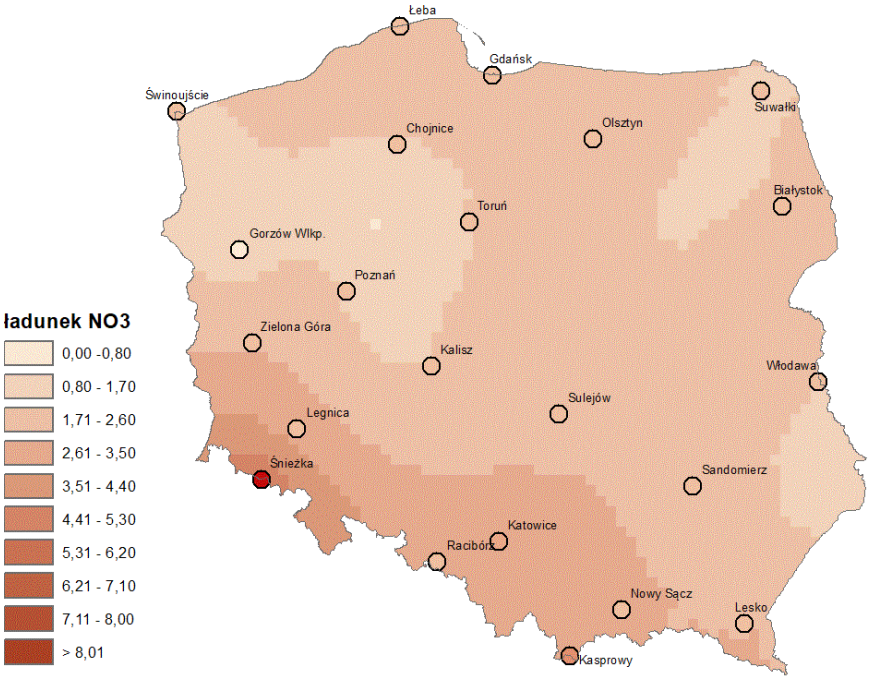


Figure 4.39 Map of NO_3^- deposition distribution obtained by combining two maps: precipitation levels and NO_3^- concentrations, both obtained by ordinary kriging with the spherical variogram function

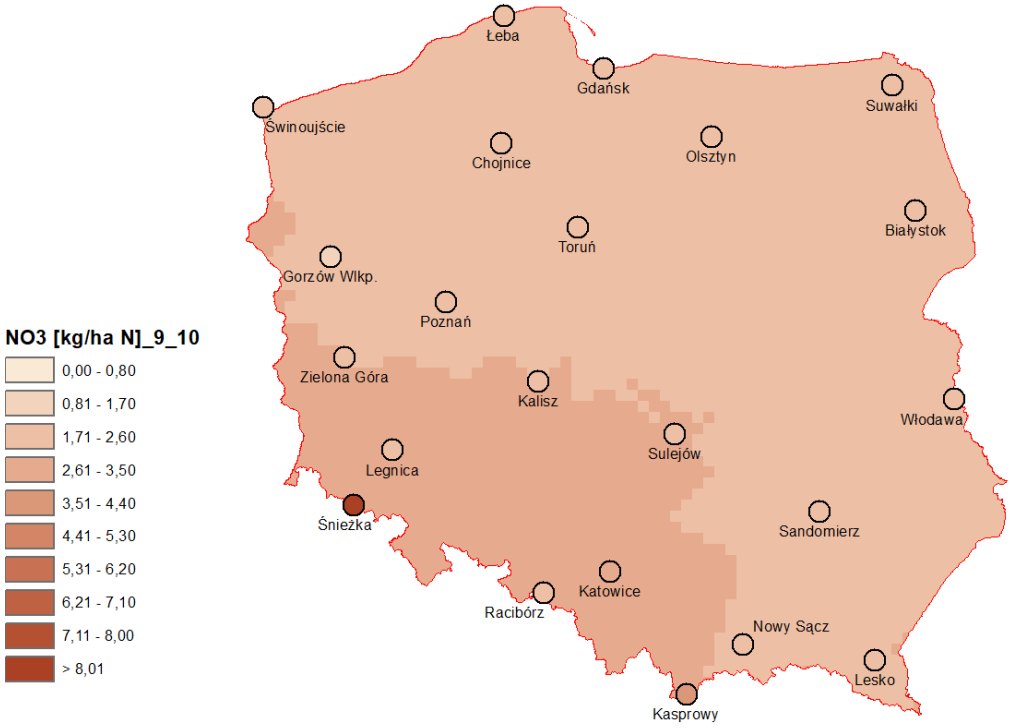


Figure 4.40 Map of NO_3^- deposition distribution obtained from data from 22 stations by ordinary kriging with the spherical variogram function

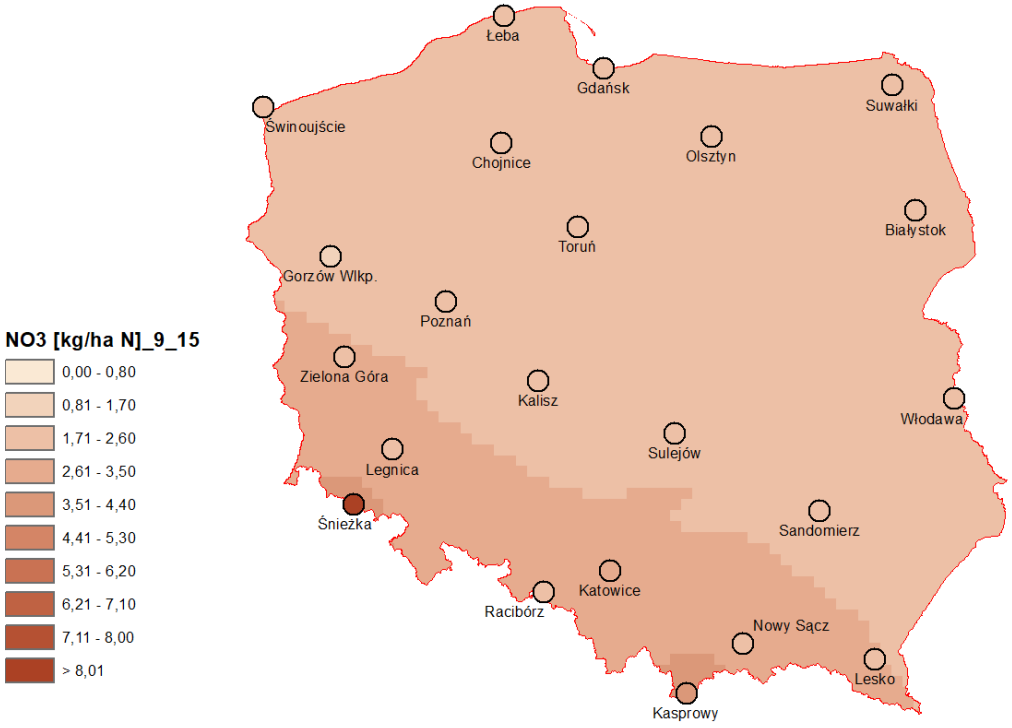


Figure 4.41 Map of NO_3^- deposition distribution obtained by cokriging (with precipitation level as second variable), obtained by ordinary kriging with the spherical variogram function

Use of data from stations in neighboring countries

The Norwegian partner NILU suggested that measurements from neighboring countries should also be used for interpolation, which should improve interpolated fields in border areas.

In accordance with the instructions provided by the Norwegian partner, the availability of data in the EEA AQ portal was checked as a first step: (https://eeadmz1-cws-wp-air02.azurewebsites.net/index.php/users-corner/statistics-e1a-table/).

For the indicators for which spatial distribution maps were prepared for this report: SO_4^{2-}, NO_3^-, NH_4^+, Cd, with reference to data from neighboring countries, it was found that only wet deposition data for cadmium in precipitation was available. However, this data is only available for German stations (84 stations, 24 of which are close to the Polish border) and for one Lithuanian station, Aukstaitija, but located quite far from the border.

A list of German stations that could be used in the spatial analysis of cadmium concentrations in precipitation in Poland is provided in the table below.

Table 4.9 EEA network stations in Germany measuring wet deposition of cadmium (based on information from EEA – European Air Quality Portal, https://eeadmz1-cws-wp-air02.azurewebsites.net/index.php/users-corner/statistics-e1a-table/)

No.	Air quality station EoI code	Air quality station Name	Air quality station Area	Latitude	Longitude	Altitude	Air quality network
1	DEBB021	Potsdam-Zentrum	urban	52.4014	13.0602	31	NET.DE_BB
2	DEBB029	Schwedt (Oder)	suburban	53.0643	14.2851	5	NET.DE_BB
3	DEBB032	Eisenhüttenstadt	suburban	52.1463	14.6382	40	NET.DE_BB
4	DEBB048	Neuruppin	suburban	52.9319	12.8095	43	NET.DE_BB
5	DEBB053	Hasenholz (Buckow)	rural	52.5639	14.0152	88	NET.DE_BB
6	DEBB064	Cottbus	urban	51.7468	14.3345	75	NET.DE_BB
7	DEBB066	Spreewald	rural	51.8975	14.0569	52	NET.DE_BB
8	DEBB067	Nauen	suburban	52.6085	12.8853	31	NET.DE_BB
9	DEBB076	Cottbus, Depo-Meisenweg	suburban	51.7761	14.3165	68	NET.DE_BB
10	DEBB083	Spremberg	suburban	51.5645	14.3764	99	NET.DE_BB
11	DEBB086	Blankenfelde-Mahlow	suburban	52.3497	13.4243	43	NET.DE_BB
12	DEBB092	Frankfurt (Oder)	suburban	52.3347	14.526	74	NET.DE_BB
13	DEBB109	Luckenwalde	urban	52.0897	13.1749	50	NET.DE_BB
14	DEBB110	Dallgow-Döberitz	suburban	52.5395	13.0488	39	NET.DE_BB
15	DEBB111	Elsterwerda	urban	51.4565	13.5225	92	NET.DE_BB
16	DESN017	Freiberg	urban	50.9163	13.3468	393	NET.DE_SN
17	DESN020	Görlitz	urban	51.1558	14.974	210	NET.DE_SN
18	DESN045	Zittau-Ost	suburban	50.8924	14.8228	230	NET.DE_SN
19	DESN051	Radebeul-Wahnsdorf	rural	51.1195	13.675	246	NET.DE_SN
20	DESN052	Zinnwald	rural	50.7315	13.7515	877	NET.DE_SN
21	DESN061	Dresden-Nord	urban	51.0649	13.7414	112	NET.DE_SN
22	DESN092	Dresden-Winckelmannstr.	urban	51.0361	13.7302	112	NET.DE_SN
23	DEUB028	Zingst	rural	54.437	12.7219	1	NET.DE_UB
24	DEUB030	Neuglobsow	rural	53.1413	13.0317	65	NET.DE_UB

The list includes 24 stations, of which only 6 are rural stations (two of them – Zingst and Neuglobsow – operate in the EMEP network and also as regional GAW WMO stations). Other locations are urban or suburban stations.

These locations provide Cd data from foreign stations, but only from the west of Poland. The source contains no data on Cd concentrations in precipitation from stations in other neighboring countries.

For other indicators, i.e. SO_4^{2-}, NO_3^-, NH_4^+, data for sulphates (precipitation), ammonium (precipitation) and nitrates (precipitation) is available for the Netherlands, Portugal and Sweden. None of these countries borders Poland, so the data cannot be used.

Another option is to use data from the EMEP network. The table below lists stations from countries bordering Poland from which data could be used.

Table 4.10 EMEP stations from neighboring countries that measure SO_4^{2-}, NO_3^-, NH_4^+ and Cd concentrations in precipitation and/or wet deposition (based on information from: EMEP/CCC-Report 1/2023 and EMEP/CCC-Report 3/2023)

No.	EMEP code	Station name	Latitude	Longitude	Altitude	Alternative name	Remarks
1	BY0004R	Vysokoe	52°20'00"N	23°26'00"E	163		excluding Cd
2	CZ0003R	Košetice (NOAK)	49°35'00"N	15°05'00"E	534	Kresin u Pacova	all
3	CZ0005R	Churanov	49°04'00"N	13°36'00"E	1118		all
4	DE0007R	Neuglobsow	53°10'00"N	13°02'00"E	62		all
5	DE0009R	Zingst	54°26'00"N	12°44'00"E	1		all
6	LT0015R	Preila	55°21'00"N	21°04'00"E	5		excluding Cd
7	SK0002R	Chopok	48°56'00"N	19°35'00"E	2008		all
8	SK0004R	Stará Lesná	49°09'00"N	20°17'00"E	808		all
9	SK0006R	Starina	49°03'00"N	22°16'00"E	345		all
PL	PL0002R	Jarczew	51°49'00"N	21°59'00"E	180		excluding Cd
PL	PL0003R	Śnieżka	50°44'00"N	15°44'00"E	1603		excluding Cd
PL	PL0004R	Łeba	54°45'00"N	17°32'00"E	2		all
PL	PL0005R	Diabla Góra	54°09'00"N	22°04'00"E	157	Puszcza Borecka	all

Nine stations are located quite close to the Polish border, but they are relatively irregularly distributed. For example, on the eastern side, only the Belarusian station Vysokoe BY04 is operational, showing the highest or near highest values in the domain for most precipitation pollutant indicators as can be seen in the attached maps from the latest 2023 EMEP reports with 2021 data.

Example maps with spatial distribution of sulphate, nitrate, ammonium and cadmium concentrations in precipitation at EMEP stations in 2021 are shown below. These maps also provide an overview of the location of measurement points in areas bordering Poland.

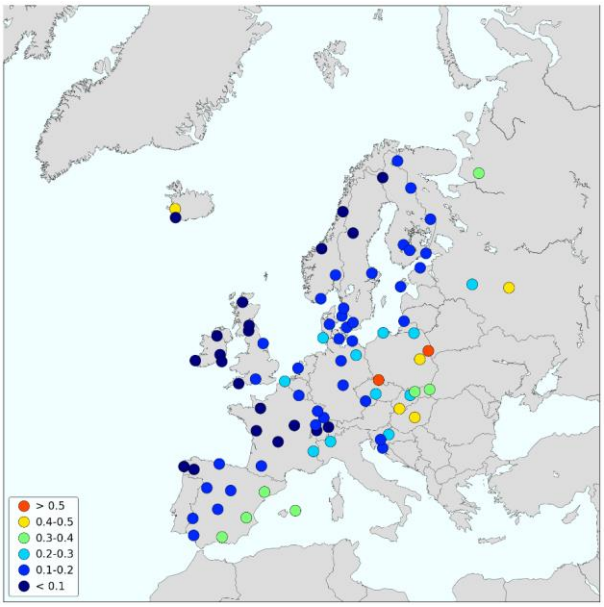
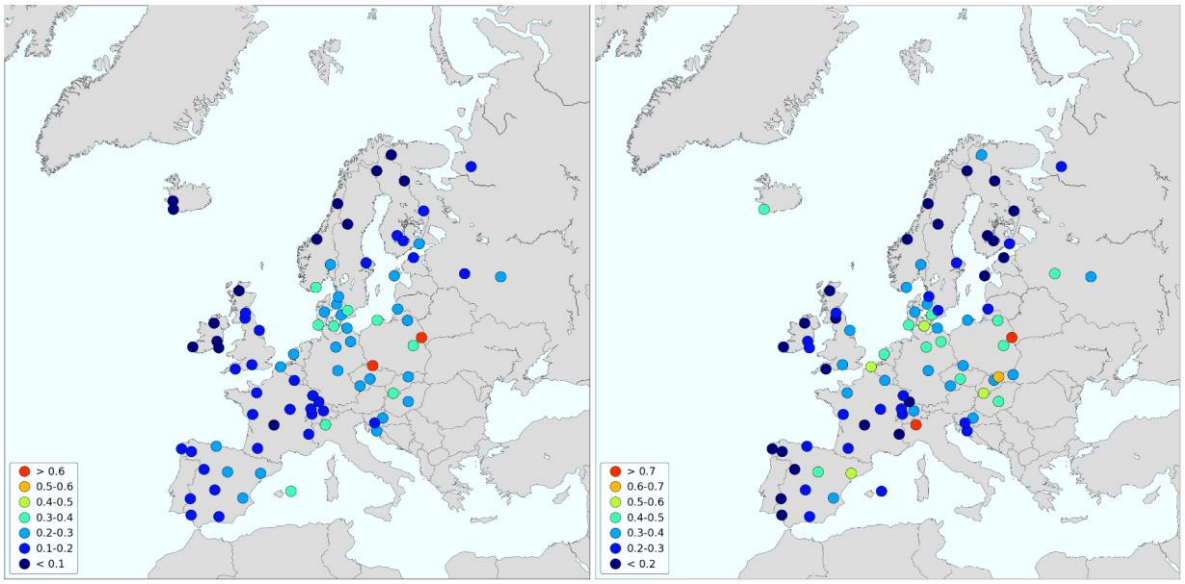


Figure 4.42 Spatial distribution of sulphate ion concentrations (values corrected for marine aerosol) [mgS/dm^3] in precipitation in 2021 [EMEP/CCC-Report 1/2023]



NO_3^- NH_4^+

Figure 4.43 Spatial distribution of nitrogen compound concentrations [mgN/dm^3] in precipitation in 2021 [EMEP/CCC-Report 1/2023]

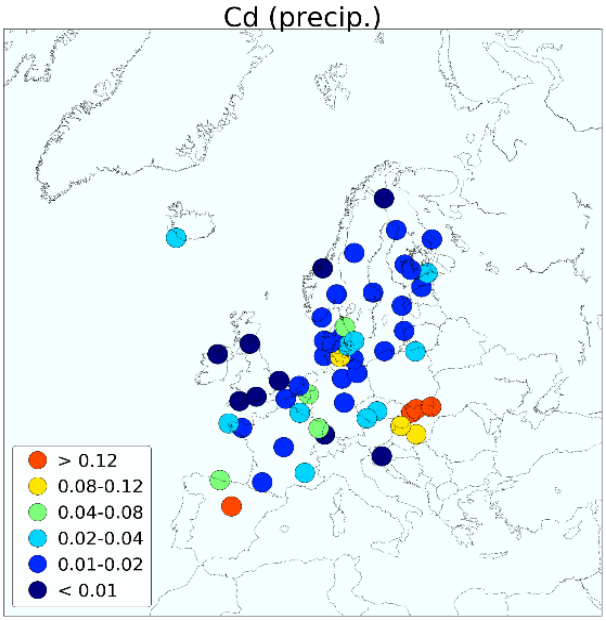


Figure 4.44 Spatial distribution of cadmium concentrations [µg/dm^3] in precipitation in 2021 [EMEP/CCC-Report 3/2023]

Unfortunately, the operational possibility of use of neighboring countries' data extracted from databases is questionable. According to the EMEP timetable, the deadline for parties to submit data to the EMEP/EBAS database is July, 31 each year for measurements made in the previous year. It usually takes about a year for the data to be verified and become available in the database. For example, for external users, at the end of Q3 2023, data from 2022 was only available from Norway and a few other countries.

The same applies to the data in the EEA database. According to the information received, in this case the deadline for countries to submit data is the end of September of the year following the year of measurement.

Currently, the assessment of atmospheric deposition in Poland for a given year is performed by the end of Q2 of the following year. Extending the time to complete the assessment
even to the end of the following year does not ensure that there will be an opportunity to use data from other databases when completing the assessment, and it is not fully justified to postpone the assessment until the following year.

An alternative solution could be to try to obtain data from Poland's neighboring countries through direct cooperation, even before the deadline for these countries to submit data to international databases. Receiving deposition data from, for example, German, Czech, Slovak or Lithuanian stations before the end of April of the year following the completion of the measurements would make it possible to use it to assess atmospheric deposition in Poland.

5. Summary

This study is final guidelines for performing atmospheric deposition assessments.

The ordinary kriging method was selected as the appropriate statistical method for atmospheric deposition assessments. This is the method currently used in Norway (Aas et al., 2022).

For homogeneous data it was proposed to use ordinary kriging without transformation and with the spherical variogram function. In the case of data involving outliers it is recommended to use ordinary kriging with logarithmic transformation and the spherical variogram function. In order to improve the accuracy of the interpolation results obtained, it is recommended to use additional data on the precipitation level and to apply the cokriging method. It was also recommended to develop maps of spatial distributions of atmospheric deposition for a grid size of 10 x 10 km.

The EMEP model was identified as the chemical transport model possible for assessing atmospheric deposition in Poland.

Also, additional possibilities for improving the quality of atmospheric deposition assessments in Poland, as suggested by the Norwegian partner, were analyzed. The analysis performed with the 2021 data showed little relevance of the model data as an aid to performing regression kriging interpolation due to the poor correlation between the measured and model data. The correlation is improved when data from high-mountain stations is removed. The new monitoring system will no longer include the current two high-mountain stations: Śnieżka and Kasprowy Wierch, but will include a new mountain station: Karkonosze. The potential future use of regression kriging and model data to interpolate atmospheric deposition requires testing on measurement data from the new monitoring network and assessing its applicability.

The tested two-step approach to calculating wet deposition distribution taking into account precipitation level data from the precipitation measurement network showed that this method is more labor-intensive and its qualitative assessment is difficult. The cokriging solution, proposed earlier, allows a more accurate assessment of the quality of the results obtained.

Taking into account the planned changes in the atmospheric precipitation chemistry monitoring system in Poland from the beginning of next year and the possibility of the availability of national data at a later date than hitherto, and assuming the acquisition of data from Poland's neighboring countries through direct cooperation, the following work timetable is proposed for the assessment of atmospheric deposition for the previous year:

- by the end of March this year – making available by GIOŚ to the assessment entity all domestic measurement data obtained in the previous year, necessary to perform the assessment for that year,
- by the end of April this year – making available by GIOŚ to the assessment entity all measurement data obtained in the previous year from stations in countries neighboring Poland, necessary to perform the assessment for that year, April – June this year – verification and preparation of the data received for statistical interpolation by the contractor,
- July – October this year – performing statistical interpolation, developing maps, performing atmospheric deposition assessment by the contractor.

6. References

Aas W., Soares J., Hamer P., Schneider P., Svendby T. and Guerreiro C. (2022). Review of methods that can be used in the assessment of atmospheric deposition. (NILU report 33/2022). Kjeller: NILU

Aas W., Halvorsen H.L., Hartz W.F., Pfaffhuber K.A., Yttri N.: Heavy metals and POP measurements, 2021, EMEP/CCC-Report 3/2023

ESRI 2001, https://www.esri.com/en-us/arcgis/products/arcgis-desktop/resources

Hjellbrekke A.-G., 2023: Data Report 2021. Particulate matter, carbonaceous and inorganic compounds, EMEP/CCC-Report 1/2023

7. List of tables

8. List of figures