Summary

Air quality is evaluated on the basis of air monitoring data. Monitoring data is often not complete enough to carry out an air quality assessment. To fill the measurement gaps, predictive models can be used, which enable the approximation of missing data. Predictive models use historical data and relationships between measured variables, including air pollutant concentrations and meteorological factors. The known predictive air quality models are not accurate, so it is important to look for models that give a lower approximation error. The use of artificial neural networks allows to reduce the prediction error compared to classical regression methods. In previous studies, a single regression model over the entire concentration range was used to approximate the concentrations of a selected pollutant. In this study, it was assumed that a group of models instead of a single model could be used for prediction. In this approach, each model from the group is dedicated to a different sub-range of the concentration of the modeled pollutant. The aim of the analysis was to check whether this approach would improve the quality of modeling. A long-term data set recorded at 2 air monitoring stations in Silesian Voivodeship (South Poland) was used in the examination. Hourly data of basic air pollutants concentrations and meteorological parameters were used to create predictive regression models. The prediction errors for the sub-range models were compared with the corresponding errors calculated for one full-range regression model. It was found that the application of sub-range models allows to reduce the modeling error of basic air pollutants.