Application of DOPPLER SODAR in short-term forecasting of PM10 concentration in the air in Krakow (Poland)

Ewa Krajny[1], Leszek Osrodka[1] and Marek Wojtylak[2]

[1] IMWM-NRI 40-045 Katowice, Bratkow 10 [2] Retired researcher (M.W.)

Correspondence to: Ewa Krajny (E.K.) ewa.krajny@imgw.pl, Leszek Osrodka (L.O.) leszek.osrodka@imgw.pl

Abstract. This paper describes an attempt to use data obtained from SODAR (Sound Detection and Ranging) for short-term forecasting of PM10 concentration levels in Krakow. Krakow is one of the most polluted cities in Central Europe (CE) in terms of PM10 concentration. This is due to the high municipal emissions. Thanks to intensive corrective actions taken by the city authorities, it is being effectively eliminated, but the unfavourable topographic location of the city limits natural ventilation. The article describes all these conditions, focusing on presenting the method of short-term correction of air quality for the time needed to take quick corrective actions by the city authorities in the event of anticipated exceedances of the permissible values. Based on several years of measurements of the physical properties of the atmosphere with SODAR, the authors of the paper suggest that SODAR data could be considered for operational use to generate short-term predictions.

The proposed method is based on the use of the spectrum, i.e. the set of amplitudes of signals returning to the SODAR receiver from the reflection of a single-frequency sound transmission and its characteristic properties depending on the physical state of the atmosphere. Similar spectra were parameterized with a single numerical value using statistical methods.

It was found that in some cases preceding high concentrations of PM10, the spectral parameters had similar values. This made it possible to develop a forecasting method for such concentrations by using data mining to search for conditions in historical data closest to the state of the atmosphere at the time of forecasting. In this part of the study, data from 2017–2018 were used. In the next step, three methods of using the SODAR data developed in this way for PM10 prediction were proposed, comparing them with the method without SODAR use. The study results were tested on independent material using data concerning the episodes of high concentrations of pollutants from October 2021 to March 2022 in Krakow. The findings were considered encouraging, also taking into account the speed and low cost of preparing the forecast.

1 Background

Air pollution is a significant problem for residents of urban agglomerations. The growing size of cities and the increasing number of inhabitants result in both an increase in emissions and a reduction in the efficient removal of pollutants outside the city limits. This then causes a periodic pollution accumulation, leading to smog episodes. There is evidence that air pollution has a significant impact on both life expectancy and quality of life. During episodes, an increased number of hospital admissions and deaths are observed, especially among children and the elderly (WHO, 2004; Bell and Davis, 2001; Dockery et al., 1993). The occurrence of high PM concentrations in Krakow is described below.

Increasing emissions as a result of lower air temperatures during the cold season, reducing the demand for heat, and thus reducing emissions as a result of "warm winters" are well-known examples of emission control mechanisms based on meteorological and climatological factors. Another and equally important role is played by thermal and dynamic conditions in the boundary layer of the atmosphere, where apart from solar radiation and dynamic factors, the degree of urbanization and land cover are also important (Xue et al., 2021; Ji et al., 2020; Xu et al., 2018; Engelbart et al., 2009; Fisher et al., 2006; Pringer et al., 2004; Arya, 1999). These issues have also been studied in Poland, where the problem of high concentrations of particulate pollutants is prevalent, especially in the winter months: January–March and October–December (Toczko, 2015; Ziemiański and Ośrodka, 2012). Among the large Polish cities, Krakow is the most polluted when it comes to PM concentration. Therefore, research in this area is particularly intensive (Bajorek and Weżyk, 2016; Bokwa, 2010; Matuszko, 2007; Walczewski, 1994). For example, extensive research on secondary sources on the legacy and contemporary research on climate diversity in Krakow can be found in Bokwa (2019) and Matuszko (2007).

The inspiration for this research was the analysis of the current state of knowledge regarding the atmospheric structure of the border layer over Krakow (Godlowska et al., 2008; Lewinska et al., 1982; Morawska-Horawska, 1978). In connection with the introduction in 1993 in Krakow of the first automatic air quality monitoring system in Poland and the delivery of system equipment, including REMTECH Doppler SODAR delivered to the Center Observation Station of the Institute of Meteorology and Water Management National Research Institute in Krakow-Czyzyny, the research became more extensive. Based on data obtained from SODAR, wind roses were developed for Krakow at the altitudes of 50 to 550 m at intervals of 100 m separately for different types of atmospheric equilibrium (Rozwoda, 1995) for mixing layers and wind circulating over the city (Fisher et al., 2006; Pronger et al., 2004) and for COST 720 (Engelbart et al., 2009), while the measurement results obtained at that time became a useful knowledge base on the subject at used in modern research (Godlowska and Kaszowski, 2019; Bajorek-Zydroń and Weżyk,

2016). In recent years, thanks to the use of drones and balloons for civilian applications, as well as miniaturization of measuring instruments, experimental studies of the vertical structure of the atmosphere from the border layer in Krakow have been widely developed, with an emphasis on the impact of meteorological conditions on the vertical profile of air pollutants (Sekula et al., 2021). These studies provide important information on the vertical distribution of PM10 concentrations under different meteorological conditions.

Intensive corrective actions were carried out in Krakow as part of research activities on the causes and effects of excessive air pollution. In Poland, legal solutions have been applied in accordance with the provisions of Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe (2008). Therefore, in areas (zones) with excessive air pollution, it is necessary to establish air quality programs to direct corrective actions (Regulation of the Minister of the Environment, 2012). In addition to an important legislative function (they constitute an act of local law), these programs are of significant preventive and educational importance. Due to the fact that the concentrations of PM10, PM2.5, and benzo(a)pyrene (B(a)P) in Krakow exceeded the limits specified in the applicable legal standards, air quality programs have been implemented for several years. The effects of their implementation include, among other things, the so-called "anti-smog resolution" (Resolution No. XVIII/243/16, 2016), which from 2019 prohibits the combustion of solid fuels for space heating and results in a gradual reduction of PM concentrations. As a result of these legislative actions, the replacement of boilers for space and water heating, a process carried out since 1995, has increased sharply, leading to a significant reduction in PM emissions in the municipal sector (the number of solid fuel boilers decreased by around 15,000 boiler installations in 2017–2019) (AQP Malopolska, 2018). This, in turn, significantly reduced PM10 concentrations, shown in Figure 1 as the number of days when the PM concentration threshold was exceeded. Specific meteorological conditions in recent years (mild and warm winters) have also contributed to the situation but have not diminished the importance of reducing municipal emissions.

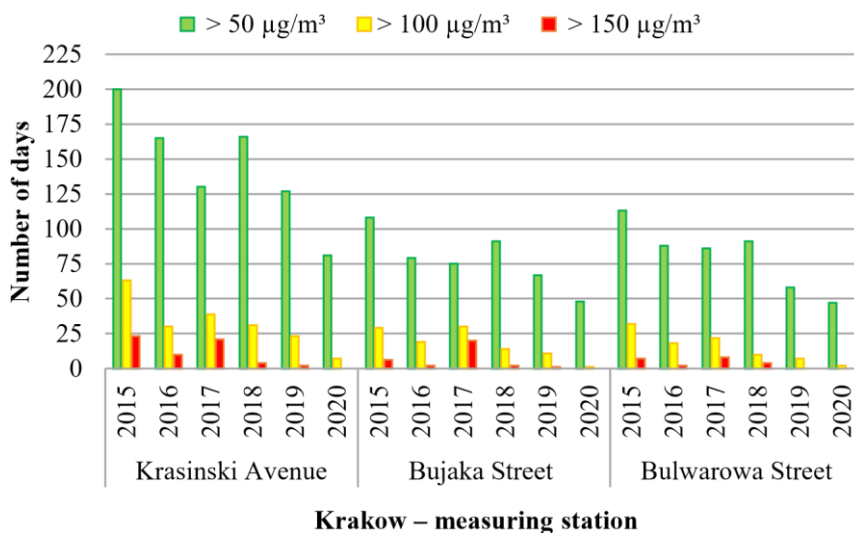



Figure 1. Total number of days with PM10 concentrations exceeding 50, 100, and 150 μg/m^3 in 2015–2020.

The unfavourable geographical location of the city also causes failure to meet air quality standards in Krakow. In the north and south, the historical centre of the city, located in the Vistula Valley, is sheltered by round hills that cross river valleys. The biggest problem of Krakow is that from the west, it is protected by the features of a topographically diverse area called the "Krakow Bridge", located at 370 m above sea level (a.s.l.) (German, 2001). The hills of the Krakow Bridge are a barrier for the entire city, reducing the efficiency of ventilation and limiting the clearance of the Vistula Valley in the western part of Krakow, threatening the influx of wind, with western air circulation dominating in Poland. The location of Krakow in the Vistula Valley affects not only the wind conditions but also contributes to temperature inversions, determining a constant type of equilibrium of the atmosphere in harsh conditions and reducing the ability to mix air vertically. Urbanization processes, which result in an increasing density of taller and taller buildings, also contribute to the reduction of wind speed and the ability to remove air pollution outside the city area. However, urbanization often contributes to improving vertical ventilation conditions by reducing the horizontal dispersion of pollutants. This is facilitated by increased anthropogenic heat production, which reduces the frequency of inversions in areas with large communities.

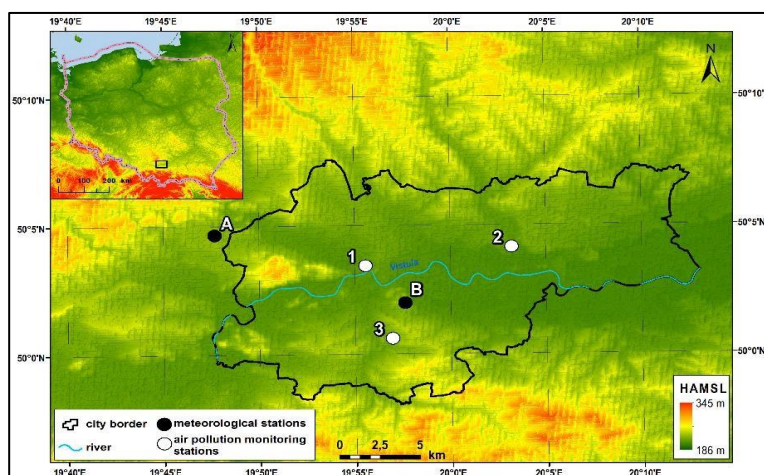



Figure 2. Geographical location of the research area showing the locations of meteorological/air pollution monitoring stations. Meteorological stations of the Institute of Meteorology and Water Management – National Research Institute (marked in black): (A) Krakow-Balice, (B) SODAR; Monitoring stations of the National Environmental Monitoring (NEM) (marked in white): (1) Krasinski Avenue, (2) Bulwarowa Street, (3) Bujaka Street.

90 ## 2 Purpose of the study

The episodes of high PM10 concentrations in Krakow prompted the city authorities to develop a system that will provide residents with access to free public transport when the limit of 8 hours of concentration is to be exceeded the next day. The implementation of an air quality forecasting system to control public transport imposes the obligation to exercise the utmost care in modelling. An inaccurate forecast is associated with the risk of the city budget incurring either unjustified costs (by multiplying the air quality forecast) or social (health) costs, which are difficult to estimate for an underestimated air quality forecast. In the specific conditions prevailing in Krakow, mainly local factors of stagnation, such situations are not uncommon. An additional element that makes it difficult to correctly forecast air quality using a deterministic model is the inability to provide an accurate emission inventory. Given this state of affairs, methods are needed to frame air quality forecasts, in particular, time trend forecasts of the absolute value of PM10 concentration. The long-term activity of Doppler SODAR in Krakow and the analysis of the impact of atmospheric

100 stability conditions on air quality gave the opportunity to use the results of these measurements to improve the air quality forecasting system (Bajorek-Zydroń and Weżyk, 2016).

3 Materials and methods

The work uses two data sources:

- Data from the selected Inspectorate for Environmental Protection/National Environmental Monitoring (IEP/NEM) (105 automatic air quality monitoring stations based in Krakow, data from 2015 to March 2022).
- Measurements of SODAR tags from 2017–March 2022.

The station characteristics and locations are shown in Table 1, and the station codes are shown in Figure 2.

Table 1. Characteristics of monitoring stations in Krakow.

Monitoring network/ Owner/ Symbol on the map	Name/Location of the monitoring station	Measured elements	Averaging time /Measurement type	GPS coordinates		Altitude m A.S.L.	Type of area/ Station type
				Latitude φ N	Length λ E		
No 1	Krasinski Ave.	PM10	1 hour Automatic	50°03′27.6″	19°55′34.3	207	urban/ communication
No 2	Bulwarowa Str.	PM10	1 hour Automatic	50°04′09,5″	20°03′12,6	195	urban/ industrial
No 3	Rocker Str.	PM10	1 hour Automatic	50°00′38,1″	19°56′57,1	223	urban/ background
IMWM-NRI A	AMS Krakow-Balice	Meteorological parameters	1 h Automatic	50°04′40″	19°47′42″	237	extra-urban (suburban)/

Monitoring network/ Owner/ Symbol on the map	Name/Location of the monitoring station	Measured elements	Averaging time /Measurement type	GPS coordinates		Altitude m A.S.L.	Type of area/ Station type
				Latitude φ N	Length λ E		
							background
Municipality of Krakow IMWM-NRIB	Swoszowicka Street SODAR (mobile station)	i.a.: vertical profile from 30 to 450 m above ground level, wind direction and speed characteristics, including atmospheric stability class DC (diffusion class)	10 min Automatic	50°02′0,63”	19°57′37,85”	247	urban/ background

110

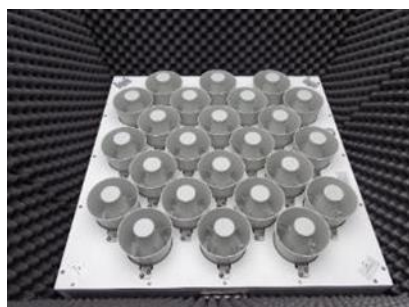
3.1 Measurements of SODAR TAGS

The monostatic Doppler SODAR PCS.2000-24 system has been operating in Krakow since January 2015 as part of the MONIT-AIR project. The system was manufactured by the German company METEK Meteorologische Messtechnik GmbH (https://metek.de). The principle of operation of SODAR is based on the Doppler effect.

The SODAR PCS.2000-24 system was built to measure wind speed and wind direction profile, based on which turbulence parameters are determined in the form of atmospheric stability class and, indirectly, air temperature inversion. It is a monostatic Doppler SODAR system transmitting antenna that transmits an audio signal of a given frequency and switches to receive and record a feedback signal. The monostatic SODAR PCS.2000-24 system transmits three audio beams simultaneously, each with three antennas: one vertical and two inclined phased antennas (Table 2).

120 The maximum range of vertical SODAR detection depends on the frequency of the transmitted signal and the current atmospheric conditions. SODAR measures the physical parameters of the atmosphere by analysing the spectrum of sound waves dispersed by fluctuations in the atmosphere of different degrees, which are the result of heat and dynamic turbulence in the atmosphere (air temperature gradients, wind speed, and hydrometeors). The Doppler effect for flat-phased antennas depends only on the wind speed and the distance between the sound transmitting the sound. The technical specifications of SODAR are shown in Table 2, and an overview of SODAR is shown in Figure 3.

Table 2. Selected technical specifications of SODAR PC.2000-24 (based on PSC.2000-24 tag manual, METEK version 2013).

Parameters	Characteristics
Operating frequency	1500...2600 Hz
Range of measured speeds of horizontal wind elements	±50 m/s
Range of measured wind directions	0...360°
Range of measured speeds of vertical wind elements	±10 m/s
Minimum working height	≥15 m (adjustable)
Minimum vertical resolution	≥5 m (adjustable)
Height range	~500 m (adjustable)

a)

b)

Figure 3. METEK SODAR PC 2000-24: a) antenna panel with phased circuit, consisting of 24 speakers; b) Swoszowicka Street in Krakow.

The methodology for compiling SODAR measurement results has been described by Emeis (2009) and Netzel et al., 130 (2000), among others, and its usefulness is demonstrated by Simakhin et al. (2017), Sgouros et al. (2011), or Yushkov et al. (2007). SODAR measurements are used for diagnostics or for setting physical property parameters in the atmospheric boundary layer (ABL). Other uses of SODAR were presented by Yang et al. (2020), who suggested using advanced SODAR data to forecast wind energy resources.

During the development of this study, SODAR was located near the Krzemionki Reservoirs at Swoszowicka Street in Krakow. It should be mentioned that these SODAR locations represent a compromise between the presentability of the measurement site for the largest possible area and the technical and organizational conditions required for the colocation of SODAR systems (Figure 2, Table 2).

SODAR measurements make it possible to determine several meteorological parameters and standards that can be directly used to analyse the air ventilation conditions in the city. These parameters include:

- wind speed (WS) and wind direction (WD) of the horizontal wind component (v and y);
- wind speed and wind direction of the vertical wind component (w);
- atmospheric stability classes (the SODAR parameter is called diffusion class (DC)), which describe the stability of the atmosphere according to the modified Pasquill-Turner classification system, expressed by letters A to F (or numbers 1 to 6), where classes A, B, C (1 to 3) mean extremely, moderately and slightly unstable atmospheric conditions; respectively, class D (4) means neutral and classes E and F (5 and 6) mean stable and highly stable atmospheric conditions.

All these elements were defined at levels up to 500 m above ground level (a.g.l.) with an approximate vertical resolution of 10 m (a.g.l.). SODAR also provides raw data that determines the transmitted and reflected signal and noise spectrum parameters. These data were taken as input for the development of a PM10 forecasting algorithm based on SODAR data.

3.2 SODAR-based PM10 prediction method

3.2.1 Justification for use

The analysis of episodes of high concentrations of PM10 in Krakow and the classes of atmospheric stability identified by SODAR led to the use of SODAR data for short-term forecasting of PM10 concentrations in the city. Figure 4 illustrates an episode of high concentrations of PM10 dust against the background of DC determined from SODAR.

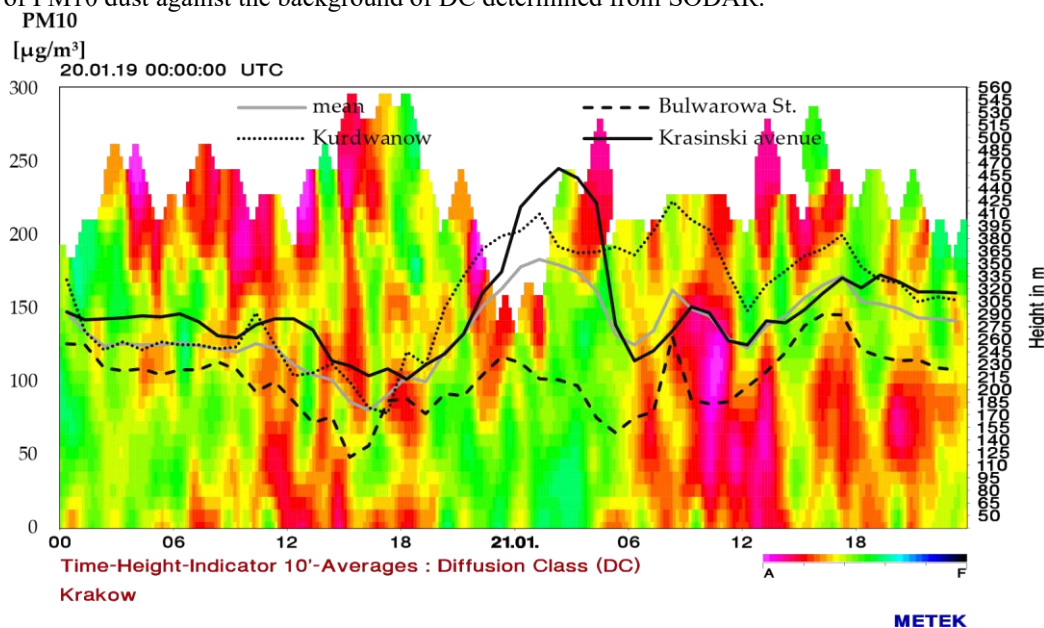



Figure 4. PM10 concentrations at selected Environmental Monitoring Stations in Krakow and the average measurements as at 20 January 2019, compared to atmospheric stability classes calculated based on SODAR data.

The lack of a clear relationship between the high DC obtained from the visualization of SODAR data and the simultaneous high concentration of PM10 dust in stagnant weather conditions (high-pressure system, low wind speed, temperature inversion conditions) suggests that the SODAR data concerning only DC is insufficient for a precise diagnosis and prediction of the episodes of high PM10 concentrations. In view of the above, it was decided to use properly processed data of the SODAR spectrum, assuming that this would allow for a more thorough analysis of ventilation conditions.

3.2.2 Preparation of SODAR data filtration

To build the PM10 concentration forecast, the basic results of SODAR measurements as a spectrum were used, i.e. a set of amplitudes of signals returning to the SODAR receiver from the reflection of sound transmission of a single frequency. Finding the required information in this way seemed to be the most promising option given the purpose of the study (which was to analyse the ventilation conditions of air in the atmosphere). When a signal is reflected through different layers of the atmosphere, it becomes scattered. The spectrum was recorded on 32 frequency channels around the transmission frequencies. In a homogeneous atmosphere, the spectra should take the form of noise. An initial transformation was required to improve the transparency of the basic data of the SODAR spectrum. After analysing the spectrum for all altitudes and on different dates, it was found that as the altitude increased, the variability of the amplitudes of all frequencies in the spectrum decreased, and a tendency to a non-zero, almost constant function appeared. In addition, secondary maxima (external sound emitters) could distort the analysis. It was also found that the shape of the spectrum was similar to the Gaussian curve. It was decided that the spectra should be simplified by subtracting their common part and similarly for spectra at all altitudes. This process is shown in Figure 5a–c. The first graph (Figure 5a) shows the original spectra measured at different altitudes, the second graph (Figure 5b) shows the spectra after subtracting their common part, and the third graph (Figure 5c) shows the common part of all spectra called the "spectral background".

The example refers to a 5-minute period on 7 October 2017. This is an illustration of the spectrum development method.

Only the filtered spectra were used for further study. In addition, in order to slightly reduce the amount of data and adjust the frequency of air volume measurements, the spectrum of interest was averaged to one hour. However, SODAR measures the data at 10-minute intervals.

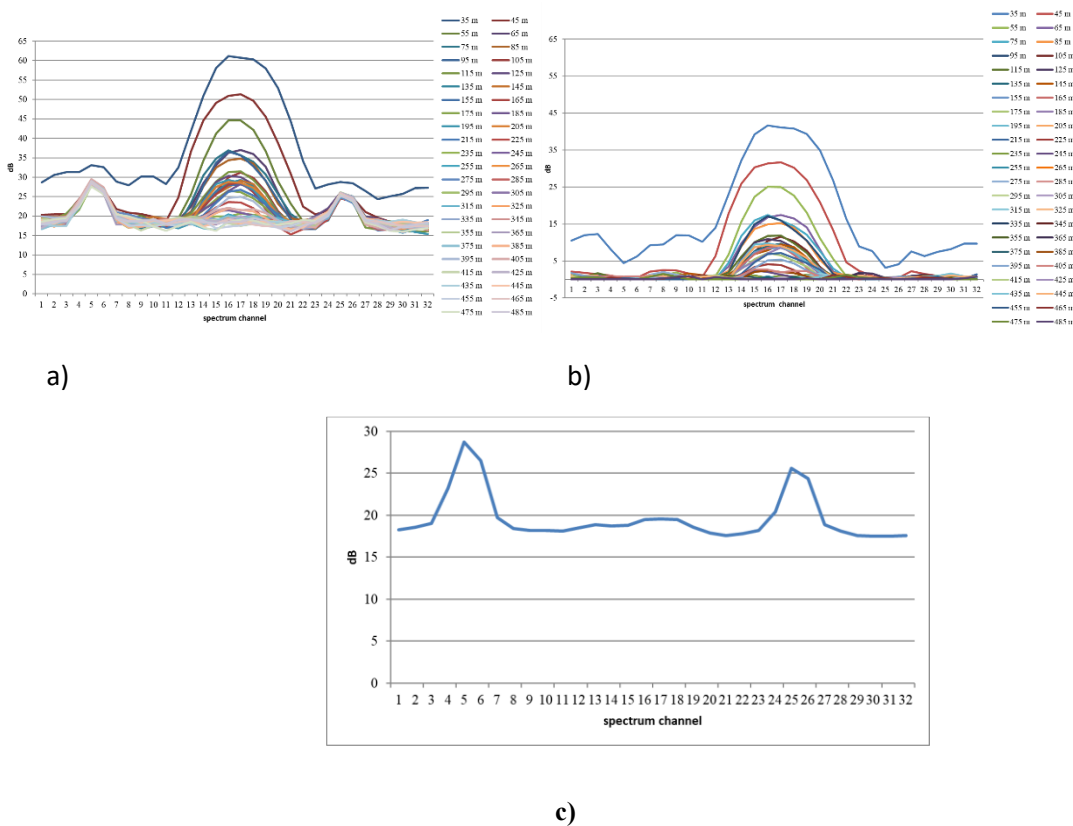



Figure 5. Example of FILTERING THE SODAR spectrum on 32 channels: a) original spectrum; b) spectrum with subtracted background; c) context.

3.2.3 Spectral properties of SODAR data

Because of the difficulty in comparing full spectra (a string of 32 values at each height) with each other and with spectra from other periods, each spectrum was characterized by a single number (parameter). Thus arose many functions of real values, the argument of which is height. Any such function will be called the "atmospheric state profile" (ASP). The most commonly observed feature of ASP is a rapid, non-linear decrease in value as the height increases. Attempts were made to determine as many parameters of the spectrum as possible. The numerical characteristics of the spectrum were chosen following a statistical approach without analysing their physical interpretations.

The characteristics with which ASP was determined were as follows:

    A.   Mean value of the SODAR reflection spectrum;

B. Maximum value of the spectra with a beam;
C. Signal-to-noise ratio (SNR);
D. Modal value (channel number with maximum high spectrum);
E. Standard deviation;
F. Median;
G. Skewness;
H. Kurtosis;
I. Similarity to the Gaussian curve.

Analysing ASP shapes, their similarity to the Φ function turned out to be familiar (Eq. 1):

$$\Phi(h) = \frac{a}{h^2+b} + c(h) \qquad (1)$$

The ASP for the parameter C profile was similar to another feature (Eq. 2):

$$\Psi(h) = a \cdot h^c \cdot e^{-b \cdot h} \qquad (2)$$

The functions Φ and Ψ were chosen arbitrarily. Parameters a, b and c of Equations (1) and (2) were determined for individual ASP (h – height above ground level using the quadratic mean approximation method). The mean-square approximation method determines the quality of the model's matching with the measurement data in the form of a coefficient of determination (the square of the correlation coefficient). This coefficient for transparency is marked RF (regularity factor) and ranges from 0 (significant irregularity) to 1 (perfect regularity).

The sample analysis suggested that RF = 0.373 means poor regularity (Figure 6a) and RF = 0.956 means good regularity (Figure 6b). A general feature was that the transformed spectra were close to zero from a certain height. Therefore, the calculations included up to 19 heights (215 m).

Figure 7 shows that the ASP at 18:00 was irregular, meaning that the vertical structure of the atmosphere was deformed and irregular. ASP at 23:00 GMT was almost perfect. This led to an investigation of the relationship between the value of ASP and PM10 concentration. For example, the chart in Figure 7 shows that 24–27 January 2018 is the period during which excessive PM10 concentrations occurred. The regularity factor for the characteristic was compared with the average concentration of PM10 from three automatic NEM stations near the SODAR location (Krasinski Avenue, Bulwarowa and Bujaka Streets). The three stations were selected for analysis because they are located in different areas of spatial development. This allowed approximating the average PM10 concentration for Krakow as they had the most complete series of measurements. For example, only one RF was shown in graphs for clarity, but the other regularity factors were similar. On this basis, it was hypothesized that ASP was disturbed a few hours before the increase in PM10 concentration. The hypothesis was developed based on the analysis of many charts, including Figure 7. This happened so often that it became the inspiration for this work.

a) b)

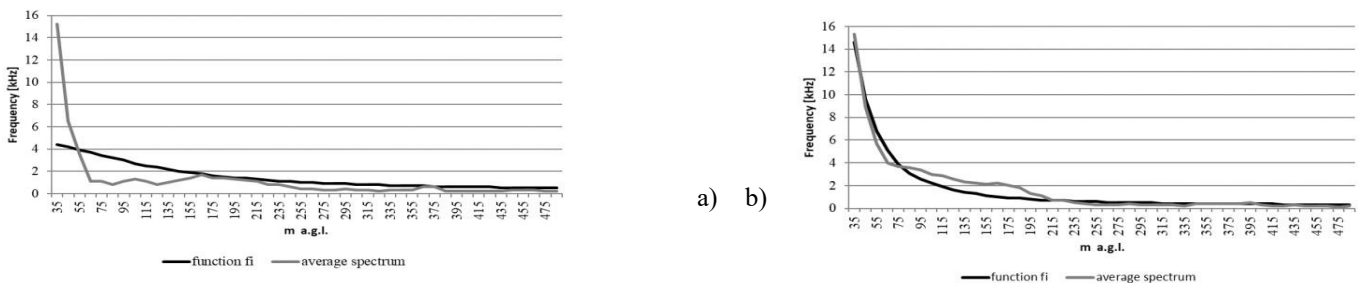



Figure 6. Graph of the regularity coefficient for parameter *A* (mean spectrum – spectrum average) compared to the PM10 concentration: a) irregularity to 18:00 UTC; b) good regularity to 23:00 UTC on 26 January 2018.

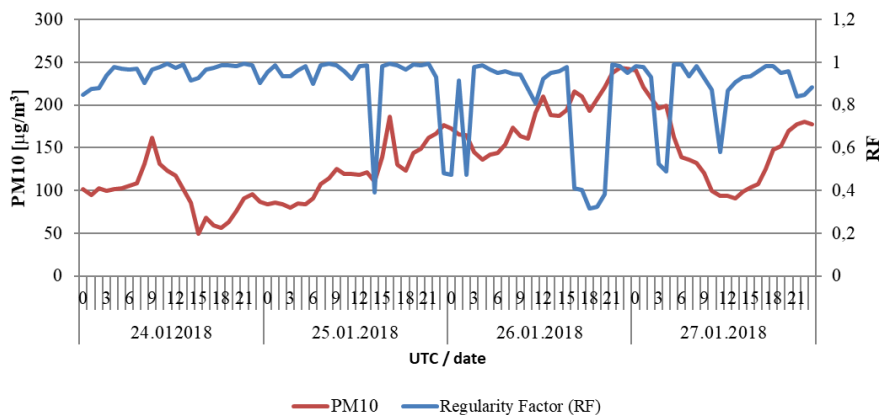



Figure 7. Example of the regularity factor (RF) for characterization (mean spectrum or spectrum average) between the SODAR data height v of 35–205 m a.g.l. and PM10 concentration.

For a complete description of the state of the atmosphere, it was decided that the meteorological RF wind characteristics should be added to the regularity coefficients A to I:

J. Horizontal wind speed (averaged in ASP);
K. Vertical wind speed (difference: min-max);
L. Wind direction (dispersion of direction around the average vector of this wind direction);
M. Temperature measured with SODAR at 2 m a.s.l.

4 PM10 forecast models

This article proposes four methods of forecasting PM10 concentration in Krakow. Three use SODAR data, and the fourth is a reference method for a forecast when only pollution data are available. The forecast is for the 12-hour time horizon, and the forecasting methods are based on SODAR and pollution data for the October–March winter seasons of 2017–2019. Reference data from 1 October 2021 to 28 February 2022 were used to compare the methods. Each method was employed to calculate the forecast every hour, and the results were compared with the actual (mathematical statistics: IEA, MAPE, MSE, Theil's UII coefficient) measurements of PM10 concentrations. Thanks to this, a quantitative assessment of each method was made, and a correlation coefficient was derived for the relationship between PM10 measurement and short-term PM10 forecast. Four methods for predicting PM10 concentrations have been developed, with descriptions and application systems of the methods explained below (Figure 8).

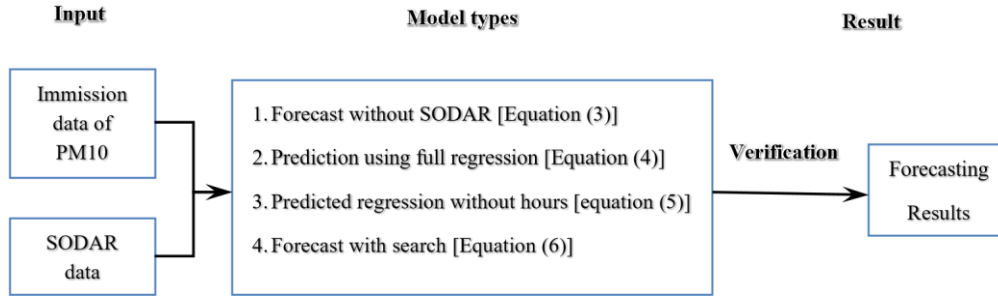



Figure 8 SODAR-based forecast generation method.

4.1 Reference method without SODAR data

Linear regression n was used to calculate the coefficients in Equation (3):

$$PM10(h + 12) = b_0 + b_1 \cos\left(\frac{\pi h}{12}\right) + b_2 \sin\left(\frac{\pi h}{12}\right) + b_3 PM10(h) \tag{3}$$

where h is the hour of the day.

4.2 Full regression method

Linear regression was used to calculate the coefficients in Equation (4):

$$\ln\big(PM10(h + 12)\big) = b_0 + b_1 \cos\left(\frac{\pi h}{12}\right) + b_2 \sin\left(\frac{\pi h}{12}\right) + b_3 \ln\big(PM10(h)\big) + \sum_{ch=A}^{M} b_{ch} RF^{ch}(h) \tag{4}$$

The corresponding transformations led to a weather pattern:

$$PM10(h + 12) = e^{b_0 + b_1 \cos\left(\frac{\pi h}{12}\right) + b_2 \sin\left(\frac{\pi h}{12}\right) + b_3 \ln(PM10(h)) + \sum_{ch=A}^{M} b_{ch} RF^{ch}(h)} \tag{4a}$$

RF^{ch} means the regression coefficient for the characteristic ch=A,…, M.

4.3 Hourless regression method

In the full regression method, the arguments (independent variables) removed factors related to daily periodicity since they also characterized SODAR factors. As a result, the following equations were formed:

$$\ln\big(PM10(h + 12)\big) = b_0 + b_1 \ln\big(PM10(h)\big) + \sum_{ch=A}^{M} b_{ch} RF^{ch}(h) \tag{5}$$

The corresponding transformations have led to a forecasting pattern:

$$PM10(h + 12) = e^{b_0 + b_1 \ln(PM10(h)) + \sum_{ch=A}^{M} b_{ch} RF^{ch}(h)} \tag{5a}$$

4.4 Search method

There is a method using an algorithm that is more flexible and can accommodate changes and improvements. The main idea of the method is to find the conditions in historical data (data mining, big data, big data form) closest to the state of the atmosphere at the time of forecasting.

Historical data and current state are expressed as a vector of numbers:

$$X(h) = [PM10(h), RF^A(h), …, RF^M(h)]^T \epsilon \mathbf{R}^{14} \tag{6}$$

The historical data for each X contain the actual values of PM10 concentration after 12 hours, which means that there is a sequence of pairs $\{(X(0), PM10(12)), (X(1), PM10(13)), …, (X(n), PM10(n + 12))\}$. Y denotes the vector of the current state. Y is related to X (0), ..., X (n), and the resulting extended sequence is normalized as follows:

$$\tilde{p} = \frac{p - min}{max - min}$$

(7)

$p - min$

Where max and min are the maximum and minimum numbers for the corresponding coordinate. The values of each coordinate are reduced to [0,1]. The sequence $\{\tilde{X}(0),...,\tilde{X}(n)\}$ is searched for the most similar vectors $\tilde{Y}$. To avoid the effects of the "dimensionality curse", which means poor differentiation of radix in Euclidean spaces greater than 10, a fractional distance is used:

$$\rho(U,V) = \sqrt[k]{\sum_{i=1}^{n}(U_i - V_i)^k} \quad U,V \epsilon \boldsymbol{R}^n$$

(8)

where $k \epsilon (0,1$

The most similar vectors are selected $X(k_1)$, ..., $X(k_s)$, and the value PM10 $(k_1 + 12)$,..., PM10 $(k_s + 12)$ is used to generate a forecast for a situation 12 h after the current state. This may be an arithmetic mean, but in practice, the geometric mean is better. The values of k and s can be specified at any time. For the purposes of this test, $k = 0.6$.

Thus:

$$r_{min} = \min \rho(Y - X(h)),$$

(9)

$h \epsilon H$

$X(h_{min})$ denotes the element that fulfils this minimum. PM10 (hmin + 12) can be taken as a 12-hour forecast in the current situation. However, such a choice is very ineffective. That is why we propose a statistical approach.

It is guided by inequality:

$$\rho(Y - X(h)) < (1 + \tau)r\_min$$

(10)

For the sample, the values of PM10 (h + 12) were selected for the periods in which the inequality is satisfied. Various descriptive statistics can be calculated based on the sample obtained this way. This may be an arithmetic mean. In practice, the geometric mean was better. In addition, good results were obtained for the median or quantiles 0.35) and 0.6). The values of τ and k can be specified at any time. For the purposes of the study, k = 0.6, τ = 0.5 was assumed. The sample size was usually around 100.

The undoubted advantage of this method is the constant supplementation of historical data. After 12 hours, when the concentration of PM10 is known, the current state (situation) is considered historical, and these latest historical data drive further forecasts. This method also uses the division of modelling results into the geometric mean and the median.

4.5 Results

Forecasting methods were verified for the entire population of areas with an average PM10 concentration in Krakow from October 2021 to March 2022 and episodes of high PM10 concentration (during which the momentary PM10 concentration was above 100 µg/m³). The initial forecast assessment compared it with a situation in which the forecast would be replaced by an average value (option 0). The first step to verify the correctness of the forecasting was to determine the forecast errors for the entire data population using the following metrics:

MAE – mean absolute error:

$$MAE = \frac{\sum_{t=1}^{T} |y_t - Y_t^*|}{T}$$

(11)

MAPE – mean absolute percentage error:

$$MAPE = \frac{\sum_{t=1}^{T} \frac{|y_t - Y_t^*|}{|y_t|}}{T} * 100\%$$

(12)

MSE – mean square error:

$$MSE = \frac{\sum_{t=1}^{T} (y_t - Y_t^*)^2}{T}$$

(13)

UII – Theil's coefficient:

$$JII = \frac{\sum_{t=1}^{T} (y_t - Y_t^*)^2}{\sum_{t=1}^{T} y_t^2}$$

(14)

Where T is the sample size (forecast length), y_t represents the measurement value, Y_t^* denotes the measurement forecast, and CORR is a correlation factor. A

Table 3. Basic statistical characteristics of the differences between forecasts and PM10 measurements for the entire data population from October 2021 to March 2022.

Statistical Parameters (# formula)	Measurement versus					
	Average Option 0	No SODAR Method (3)	Full Regression Method (4)	Regression No hours Method (5)	Search on Geometric Mean Method (6)	Median search
MAE (11)	20.8	15.87	16.09	16.15	15.76	16.16
MAPE (12)	97.1	65	67.2	68.2	67.9	69.4
MSE (13)	727	477	474	475	445	471
UII (14)	0.326	0.214	0.212	0.213	0.199	0.211
CORR	0	0.591	0.657	0.648	0.636	0.626

The results concerning the whole population revealed that the error results were better for all forecasting methods than if the average value was used (Table 3). This means that the forecasts have been made as intended. Analysis of specific model data revealed various statistics on forecast errors, while the MAE (11) achieved the lowest values for the prediction of the search method using a geometric mean (15.76). MAPE (12) had the lowest value for the reference method from the no SODAR data (65%). MSE (13) and UII (14) proved that the search method using geometric mean was the best. CORR analysis showed that it was best with the full regression method. The results of such analyses for the whole population did not show a preference for a single forecasting method.

In the second stage, episodes of high PM10 concentration were distinguished for detailed analysis. Episodes of high PM10 concentration are critical for providing free public transport in Krakow (where PM10 > 100 μg/m³). Four such episodes were recorded from October 2021 to March 2022: 12–16 December 2021, 25–30 December 2021, 23–26 January 2022, 13–16 February 2022. The trend in the observed PM10 concentrations and the projections based on four interesting methods are shown in Figures 9 to 12.

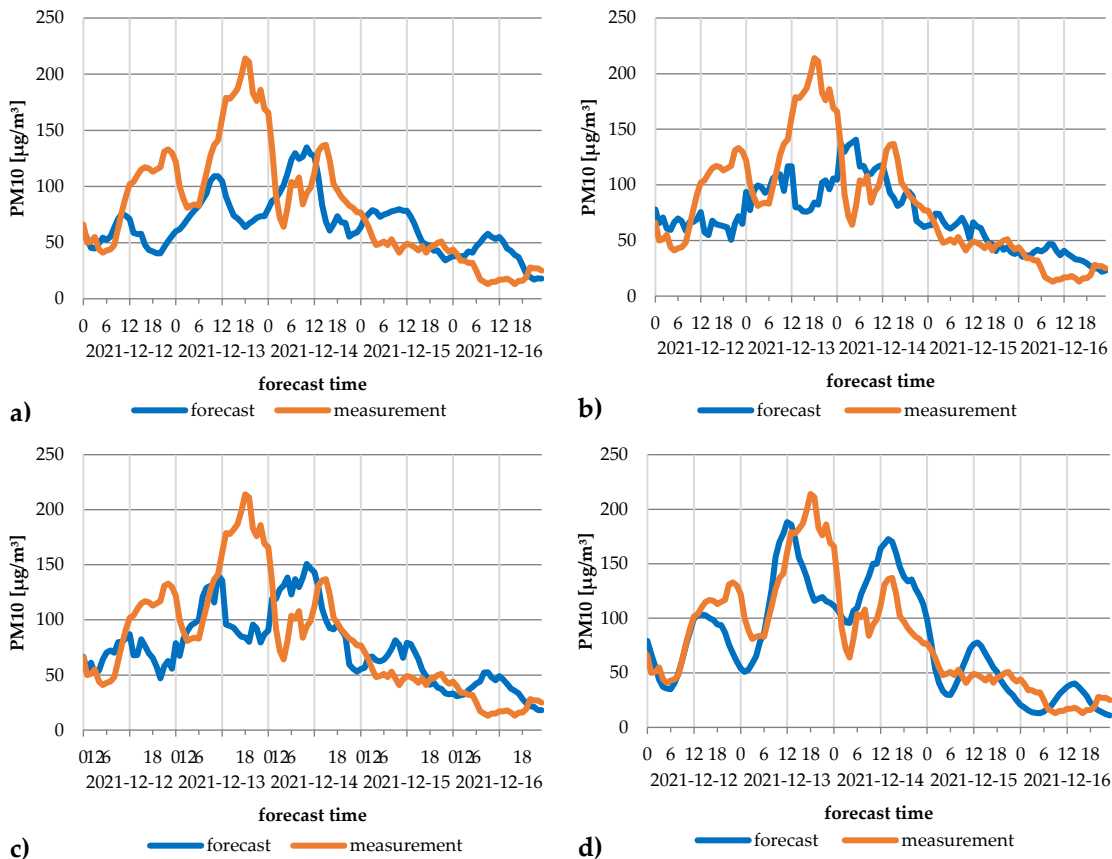



Figure 9. Measurements and forecasts for 12–16 December 2021: a) forecast without SODAR data; b) forecast predicted by regression without hours; c) full regression forecast; d) search method forecast.

Figure 10. Measurements and forecasts for 25–30 December 2021: a) forecast without SODAR data; b) forecast predicted by regression without hours; c) full regression forecast; d) search method forecast.

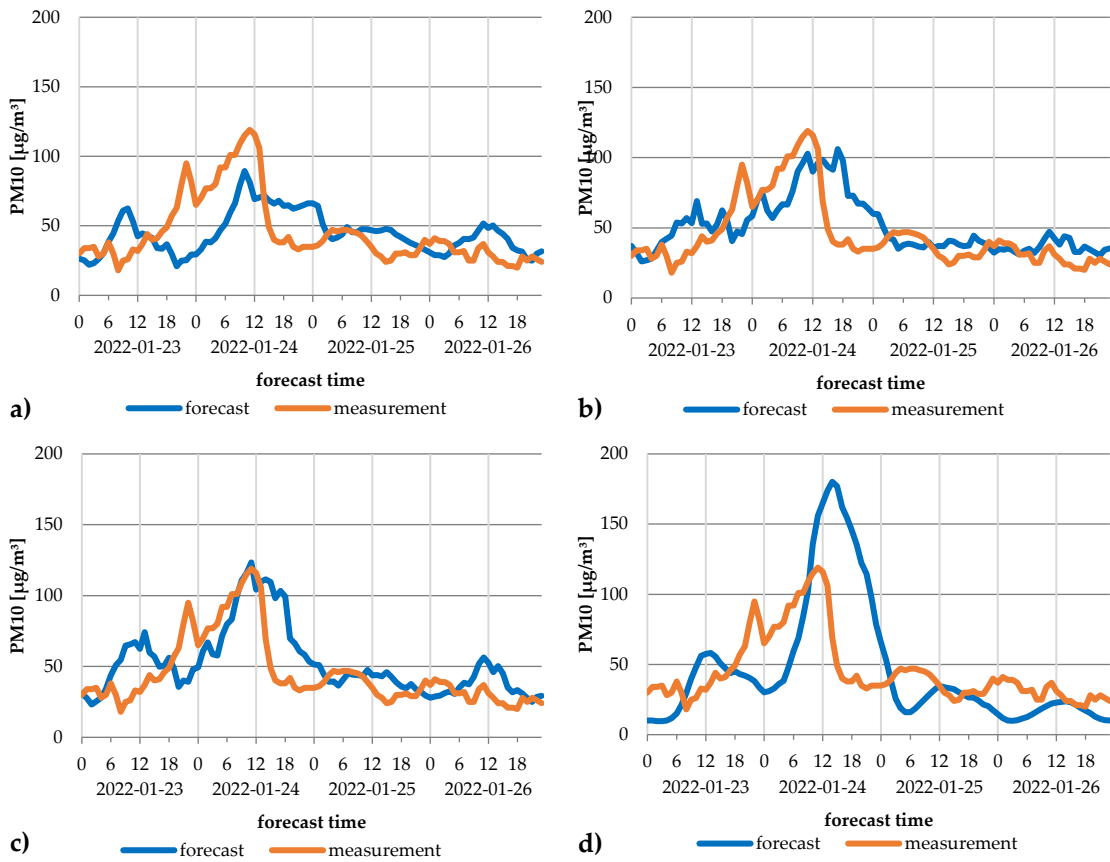



Figure 11. Measurements and forecasts for 23–26 January 2022: a) forecast without SODAR data; b) forecast predicted by regression without hours; c) full regression forecast; d) search method forecast.

Figure 12. Measurements and forecasts for 13–16 February 2022: a) forecast without SODAR data; b) forecast predicted by regression without hours; c) full regression forecast; d) search method forecast.

The above trends in PM10 predictions compared to measurements during PM10 episodes > 100 μg/m³ (Figures 9 to 12) revealed that each prediction method underestimated the measured maximum PM10 concentration. It should also be noted that the discrepancy between predictions and measurements changed with episodes. It can be concluded that the meteorological origin of each episode was different, which made forecasting difficult. Despite these discrepancies, it was found that each forecasting method provided a trend result that mimicked the measurements, sometimes with a delay of several hours. The methods fit well under certain circumstances. This was the case, for example, for the forecast (d) for the section 12–16 December 2021 (Figure 9d), forecast (d) for the section 25–30 December 2021 (Figure 10d) and the forecast (c) for the section from 23–26 January 2022 (Figure 11c).

370

Table 4. Basic statistical characteristics of the differences between PM10 projections and measurements for 100 PM10 episodes μg/m³ from October 2021 to March 2022.

Statistical Parameters (# formula)	Measurement versus					
	Average Option 0	No SODAR Method (3)	Full Regression Method (4)	Regression No hours Method (5)	Search on Geometric Mean Method (6)	Median search
MAE (11)	33.35	26.61	23.46	22.95	24.16	23.89
MAPE (12)	74.1	70.7	62.9	60.8	60.4	61.6
MSE (13)	2207	1343	1013	1010	1141	1113
UII (14)	0.397	0.242	0.182	0.182	0.205	0.2
CORR	0	0.475	0.62	0.627	0.599	0.593

The aggregated characterization of the forecast error statistics for episodes (Table 4) revealed that, as with the entire data population, each forecasting method produced better results than the average values, meaning the predictions were useful. As for the statistical parameters studied, the fifth method – regression without hours (with the lowest MAE, UII and MSE and the highest CORR) was the best fit. However, this does not mean this method is universal or suitable for every episode, as shown above.

5 Conclusions

The use of PM10 forecasts for short-term improvement of air quality is becoming more and more frequent in Poland (three-day forecasts and forecasts of the Chief Inspectorate for Environmental Protection, available online: http://powietrze.gios.gov.pl/pjp/airPollution). However, air quality forecasting is rarely used to guide administrative and economic decision-making (e.g. providing free public transport). This is because inaccurate forecasts cause high social costs (dissatisfaction of residents) or unjustified financial costs (lack of revenue from public transport tickets). Therefore, applying air quality models to these forecasts must ensure as little loss as possible due to poor decision-making.

The research discussed here shows that using SODAR data to support an air quality forecasting system is reasonable. In particular, the following proposals were made:

- The SODAR model can be complementary to other forecasting methods, as it is highly useful due to its simplicity and speed of calculations.
- The SODAR model does not require emission data, for which temporal and spatial variability are difficult to verify quickly.
- Table 4 shows that, especially at high concentrations, SODAR data provide significant information relative to the model (3) without SODAR.
- The use of simple formulas for regression models in forecasting while maintaining their multivariance (taking into account the four forecast options) facilitates the optimization of the predictive process.
- The model is ready for use, but work is underway to improve it through a different selection of SODAR parameters.

Applying the model proposed in this article may improve short-term air quality predictions, although the model still requires further testing, especially for episodes of high PM10 concentrations.

Author Contributions: idea and conception, E.K. and L.O.; methodology, E.K.; software, M.W.; validation, M.W. and E.K.; formal analysis, L.O.; investigation, L.O.; resources, E.K.; data curation, E.K.; writing – original draft preparation, L.O.; writing – review and editing, E.K.; visualization, E.K.; supervision, M.W.; project administration, L.O.; funding acquisition, L.O. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was also developed from the research results produced by the AIR BORDER project, funded by the Ministry of Science and Higher Education of the Republic of Poland, Contract No. 3911/INTERREG VA/2018/2. The research was co-financed from the science fund under INTERREG CE AIR TRITIA, and by the Ministry of Science and Higher Education of the Republic of Poland (Contract No. CE1101). The MONIT-AIR project was co-financed with funds from the European Economic Area's 2009-2014 Financial Mechanism.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The meteorological data used in this paper (the wind roses) are available to the public: https://danepubliczne.imgw.pl/ (accessed July 2022). The air quality data used in this paper are available to the public: https://powietrze.gios.gov.pl/pjp/archives/ (accessed July 2022). The owner of Sodar is the Krakow City Hall, but it is lent to IMWM-PIB. The measurement results are not made publicly available because they constitute raw data and are prepared periodically after validation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

AQP Air Quality Portal. Forecast per day. Chief Inspectorate for Environmental Protection. Available online: https://powietrze.gios.gov.pl/pjp/airPollution/ (accessed January 2022).

Arya, P.S. Air Pollution Meteorology and Dispersion. Oxford University Press: Oxford, UK, 1999.

Bajorek-Zydroń K. and Węzyk, P. Atlas Pokrycia Teren i Naa Aeration of the City of Krakow; Krakow City Hall, Department of Environmental Management: Kraków, Poland, 2016. Available online: http://geo.ur.krakow.pl/download/pobierz.php?file=publikacje/literatura/Wezyk_Atlas_2016_tekst.pdf (access December 2018).

Bell, M.L.; Davis, D.L. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. Environ. Health Perspect. 2001, 109 (Suppl S3), 389–394. https://doi.org/10.1289/ehp.01109s3389.

Bokwa, A. Development of research on the local climate of Krakow. Acta Geogr. Ice. 2019, 108, 7–20. https://doi.org/10.26485/AGL/2019/108/1.

Bokwa, A. Many years of changes in the s tructure of the mezoklimat city on the example of Krakow; Institute of Geography and Spatial Management of the Jagiellonian University: Kraków, Poland, 2010.

Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on air quality and cleaner air for

Europe (Journal of Laws EU.L.2015.226.4). Available online: L_2008152EN.01000101.xml (europa.eu) (accessed on 10 July 2022)

Dockery, D.W.; Pope, C.A.; Xu, X.; Spengler, J.D.; Ware, J.H.; Fay, M.E.; Ferris Jr, B.G.; Speizer, F.E. An association between air pollution and mortality in six US cities. N. Engl. J. Med. 1993, 329, 1753–1759.

Emeis, S. Surface-Based Remote Sensing of the Atmospheric Boundary Layer; Springer: Dordrecht, Netherlands; Heidelberg, Germany; London, UK; New York, NY, USA, 2009. https://doi.org/10.1007/978-90-481-9340-0

Engelbart, D.Monna, Wim & Nash, John & Matzler, Christian. (2009). EUR 24172 - COST-Action 720 - Integrated Ground-Based Remote-Sensing Stations for Atmospheric Profiling. Final Report of COST Action 720

Fisher, B.; Kukkonen, J.; Piringer, M.; Rotach, M.W.; Schatzmann, M. Meteorology applied to urban air pollution problems: Concepts from COST 715. Atmos. Chem. Phys. Discuss. 2006, 6, 555–564. https://doi.org/10.5194/acp-6-555-2006 .(access Janury 2007).

German, K. F Physico-geographical Regions of the Lesser Poland Voivodeship. Folia Geographica. Series: GeographicaOeconomica, 2001; Volume 31, p.32. Available online: https://rcin.org.pl/igipz/dlibra/showcontent/publication/edition/65112?showContent=true&id=65112 (accessed July 2021).

Godłowska, J.; Kaszowski, W. Testing various morphometric methods for determining the vertical profile of wind speed above Krakow, Poland. Bound.-Layer Meteorol. 2019, 172, 107–132. https://doi.org/10.1007/s10546-019-00440-9.

Godłowska, J.; Tomaszewska, A.M.; Hajto, M. Relation of the Amount of Pollution Immission in Krakow with the Conditions in the Municipal Border Layer, Determined on the Basis of Sodar Data. In Climate and Bioclimate of Cities.

Kłysik, K.J., Wibig, K., Fortuniak, W., Eds.; Wydawnictwo Uniwersytetu Łódzkiego: Łódź, Poland, 2008; pp. 455–465. Implementation of the Air Quality Plan for the Malopolska Region – Malopolska in a healthy atmosphere". 2018. Available online: https://powietrze.malopolska.pl/wp-content/uploads/2021/07/AQP_Malopolska_Full_text.pdf (accessed on 11 July 2021).

Ji, M.; Jiang, Y.; Han, X.; Liu, L.; Xu, X.; Qiao, Z.; Sun, W. Spatiotemporal Relationships between Air Quality and Multiple Meteorological Parameters in 221 Chinese Cities. Complexity 2020, 2020, 1–25. https://doi.org/10.1155/2020/6829142.

Lewińska, J.; Bartosik, J.; Baścik, J.; Red, M.; Zgud, K. The Influence of the City on the Local Climate (on the example of the Krakow Agglomeration); Institute of Environmental Management: Warsaw, Poland, 1982; pp. 1–162.

Matuszko, D. Ed. The climate of Krakow in the twentieth century; IGiGP UJ: Kraków, Poland, 2007; p. 250.

Morawska-Horawska, M. The course of variability of the vertical wind speed profile in the 300 m layer in Krakow. IMGW News: Warsaw, Poland, 1978; Volume 25, pp. 123–129.

Netzel, P.; Stano, S.; Zarebski, M. Three-monostatic Doppler sodar 3DDS. IMGW News 2000; Volume 23, pp. 131–136. Available online: https://www.worldcat.org/title/trojmonostatyczny-sodar-dopplerowski-3dds/oclc/998387930 (accessed 20 July 2022).

Piringer, Martin & Joffre, Sylvain & Baklanov, Alexander & Burzynski, Jerzy & De Ridder, Koen & Deserti, Marco & 440 Karppinen, Ari & Mestayer, Patrice & Middleton, Douglas & Tombrou, Maria & Vogt, Roland & Christen, Andreas. (2004). The urban surface energy budget and the mixing height: Some results of recent European experiments stimulated by the COST – Action 715.

Regulation of the Minister of the Environment of 11 September 2012 on Air Quality Programmes and Short-Term Action Plans (Journal of Laws of 2012, item,1028). Available online: http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20120001028 (accessed on 10 July 2022).

Resolution No. XVIII/243/16 of the Małopolska Regional Assembly of 15 January 2016 on the introduction of Restrictions on the operation of Installations in which Fuel Spillage takes place in the Area of the Municipality of Kraków. Available online: http://edziennik.malopolska.uw.gov.pl/legalact/2016/812/ (accessed on 10 July 2022).

Rozwoda W. Altitude characteristics of wind in Krakow depending on the equilibrium states of the atmosphere, Wiadomości IMGW, 1995, 18 (3-4), 109-117.

Sekuła, P.; Bokwa, A.; Bartyzel, J.; Bochenek, B.; Chmura, Ł.; Gałkowski, M.; Zimnoch, M. Measurement report: Effect of wind shear on PM10 concentration vertical structure in the urban boundary layer in a complex terrain. Atmos. Chem. Phys. Discuss. 2021, 21, 12113–12139. https://doi.org/10.5194/acp-21-12113-2021.

Sgouros, G.; Helmis, C.G.; Degleris, J. Development and application of an algorithm for the estimation of mixing height with the use of a SODAR-RASS remote sensing system. Int. J. Remote. Sens. 2011, 32, 7297–7313. https://doi.org/10.1080/01431161.2010.521205.

Simakhin, V.A.; Cherepanov, O.S.; Shamanaeva, L.G. November. Robust Statistical Methods in Sodar Studies of the ABL. In 23rd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics; International Society for Optics and Photonics: Bellingham, WA, USA, 2017, Volume 10466, p. 1046632.

Struzewska, J. and Kaminski, J.W., 2012. Impact of urban parameterization on high resolution air quality forecast with the GEM–AQ model. Atmospheric Chemistry and Physics, 12(21), pp. 10387-10404.

Toczko, B. Fine dusts in the Atmosphere; PIOŚ: Warszawa, Poland, 2015.

Walczewski, J., Ed. Characteristics of the Border Layer of the Atmosphere Above the City [on the example of Krakow]. Research Materials. Series: Meteorology; IMGW: Warsaw, Poland, 1994, volume 22, p. 109.

World Health Organization. Regional Office for Europe. (2004). Health aspects of air pollution : results from the WHO project "Systematic review of health aspects of air pollution in Europe". Copenhagen : WHO Regional Office for Europe

https://apps.who.int/iris/handle/10665/107571Availablehttps://apps.who.int/iris/bitstream/handle/10665/107571/E83080.pdf?seq uence=1&isAllowed=y (accessed on 12 December 2021.

Xu, Y.; Xue, W.; Lei, Y.; Zhao, Y.; Cheng, S.; Ren, Z.; Huang, Q. Impact of Meteorological Conditions on PM2.5 Pollution in China during Winter. Atmosphere 2018, 9(11), 429 https://www.mdpi.com/2073-4433/9/11/429 (accessed January 2019).

Xue, W.; Shi, X.; Yan, G.; Wang, J.; Xu, Y.; Tang, Q.; Wang, Y.; Zheng, Y.; Lei, Y. Impacts of meteorology and emission variations on the heavy air pollution episode in North China around the 2020 Spring Festival. Sci. China Earth Sci. 2021, 64, 329–339. https://doi.org/10.1007/s11430-020-9683-8.

Yang, G.; Xinlei, S.; Baoliang, L.; Wenzhong, S.; Mingjiang, Z.; Ziyan, Z. November. Research on wind power prediction based on doppler sodar. In 2020 Chinese Automation Congress (CAC); IEEE: Piscataway, NJ, USA, 2020; pp. 1345–1348. Yushkov, V.P.; Kallistratova, M.A.; Kouznetsov, R.D.; Kurbatov, G.A.; Kramar, V.F. Experience in measuring the windvelocity profile in an urban environment with a Doppler sodar. Izv. Atmos. Ocean. Phys. 2007, 43, 168–180. https://doi.org/10.1134/s0001433807020041.

Ziemiański, M. and Ośródka, L., Eds. Climate Change and Monitoring and Forecasting the State of the Atmospheric Environment. Institute of Meteorology and Water Management National Research Institute: Warsaw, Poland, 2012.

500