

Jakub M. GAC<sup>1\*</sup> and Monika PETELCZYC<sup>2</sup>

## RECONSTRUCTION OF DYNAMICS OF SO<sub>2</sub> CONCENTRATION IN TROPOSPHERE BASED ON RESULTS OF DIRECT MEASUREMENTS

### REKONSTRUKCJA DYNAMIKI STĘŻENIA SO<sub>2</sub> W TROPOSPFERZE NA PODSTAWIE WYNIKÓW BEZPOŚREDNICH POMIARÓW

**Abstract:** A method for the reconstruction of the dynamics of processes with discrete time, developed in our previous papers, has been applied for study the dynamics of concentration of sulfur dioxide in lower troposphere. For the analysis, recordings of sulfur dioxide concentration from four measurement stations located in Poland (two of them has been located in huge cities and two in rarely inhabited regions) were used. We managed to obtain the deterministic and stochastic component of this dynamics. In result, we estimate the lifetime of sulfur dioxide in troposphere and the increase of sulfur dioxide concentration influenced by anthropogenic sources.

**Keywords:** air pollution, SO<sub>2</sub> emission, time series, Langevin equation, denoising

## Introduction

Air pollution due to industrial activity of people, is a serious problem for people health and environment. The air pollution is a situation, in which substances that occur from anthropogenic activities or natural sources are present at concentrations sufficiently high above their normal ambient levels. Exceeded levels of concentration causes a measurable effect on humans or their environment (animals, vegetation, materials) [1]. These substances, which may be present in form of gases, liquids (small droplets) or solids (micro-size particles) are usually composed of four main groups of substances. Pollutants are simple inorganic gases (sulfur dioxide, nitrogen oxides, ozone etc.), organic substances (starting from simple hydrocarbons up to very complex multifunctional macromolecules), heavy metals and their compounds (e.g. lead and its oxides, mercury) and the other substances such as soot particles. The great importance of air pollutants has been a subject of intensive studies for decades, especially in contests of their influence of human health [2-4]. Other explorations are dedicated to processes of neutralization of harmful

---

<sup>1</sup> Faculty of Chemical and Process Engineering, Warsaw University of Technology, ul. L. Waryńskiego 1, 00-645 Warszawa, Poland, phone +48 22 234 65 08

<sup>2</sup> Faculty of Physics, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland, email: Monika.Petelczyc@pw.edu.pl

\*Corresponding author: Jakub.Gac@pw.edu.pl

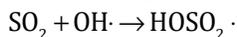
components as well as to limitations of their emissions [1]. Finally, knowledge about natural transformations of these substances in atmosphere may be useful for estimation their lifetime. For last purpose, the direct measurements of concentrations of considered substances are necessary. However, to find the optimal method for limitation of the negative influence of these substances, the knowledge about their physical and chemical transformations is needed. There are many experimental works investigating various paths of chemical reactions. Knowledge about kinetics of reactions cannot be gathered directly from measurements, while they are influenced not only by emissions and by reactions taking place in atmosphere but also by some stochastic or nearly stochastic processes, e.g. air circulation (winds) or humidity changes.

Thus in the paper, we propose the analysis of recorded concentrations of pollutant gases by stochastic approach. Proposed method extracts deterministic part of the signal and its additive noisy component [5]. Deterministic part is an internal mechanism, which mainly governs the observed process and can be a quantitative descriptor of its origin. The noise component is responsible for external (but even natural) perturbations, which have impact on resultant fluctuations of analyzed phenomenon. These two elements are reconstructed directly from time series in form of polynomial functions and build discrete model of time evolution of the process.

In present work, this method is applied to sulfur dioxide ( $\text{SO}_2$ ) concentration dynamics. The choice of  $\text{SO}_2$  arises from the fact, that this gas is one of most important anthropogenic gaseous pollution. On the other hand, there are a reach database of measurements of its concentration - both in big cities as well as outside of them. Concentration of  $\text{SO}_2$  in various regions and its dynamics has been a subject of interests of many investigators in the last years [6-8].

Global emission of sulfur dioxide is estimated as equal to 64.2 Tg of sulfur per year [9] what makes  $\text{SO}_2$  the main sulfur-containing compound of pollutant gases. The largest part of this emission is caused by fossil fuel combustions and by industry sources (56.3 Tg of sulfur per year). Another anthropogenic source of  $\text{SO}_2$  is biomass burning (ca. 1.3 Tg of sulfur per year). Sulfur dioxide is also present in volcanic gases. In summary, the highest concentration of  $\text{SO}_2$  and the highest emission is observed in big cities and huge industrial regions.

Gaseous sulfur compounds emitted from mentioned sources undergoes a few physical or chemical transformations which are responsible for decrease of its concentration. Sulfur dioxide reacts under tropospheric conditions via both gas- and aqueous-phase processes [1]. For gas phase processes, the most important reaction is a process with hydroxyl radicals, which are common byproducts of various atmospheric reactions [10]:



$\text{HOSO}_2\cdot$  transforms further to sulfur trioxide  $\text{SO}_3$  and sulfuric acid  $\text{H}_2\text{SO}_4$ . Aqueous-phase processes usually start with dissolution of  $\text{SO}_2$  in water and its dissociation. Dissolved  $\text{SO}_2$ , bisulfite and sulfite ions may then undergo reactions in liquid phase with ozone [11, 12], hydrogen peroxide [13] and organic peroxides [14, 15]. There is also possible oxidation of  $\text{SO}_2$  with oxygen, catalyzed by ferrum or manganese ions [16].

The main physical process leading to decrease of  $\text{SO}_2$  concentration is a dry deposition [17].

Most of these reactions are of first order, thus we expect, that the dynamics of  $\text{SO}_2$  concentration in atmosphere is given by:

$$x(t + \Delta t) = x(t) - kx(t)\Delta t + E\Delta t + g[x(t)]\xi \quad (1)$$

where  $x(t)$  denotes the concentration of SO<sub>2</sub> in time  $t$ ,  $E$  is a measure of mean emission of SO<sub>2</sub> in the vicinity of measurement station (the precise definition: hourly increase of SO<sub>2</sub> concentration in air, caused by SO<sub>2</sub> sources),  $k$  denotes the total rate of decay of SO<sub>2</sub> and  $\Delta t$  is time period between two successive measurements (in our case it is an hour). Last term in eq. (1) represents the stochastic term, which influence on the concentration of SO<sub>2</sub> by atmospheric circulation (natural phenomena), short-period changes of emission level etc. We expect that this stochastic term is a product of a function  $g$ , which may depend on SO<sub>2</sub> concentration  $x(t)$ . Stochastic variable  $\xi$  has a distribution with zero mean and unit variance.

The main goal of this paper is to describe the kinetics of sulfur dioxide decay basing on denoising algorithm. Especially, the lifetime of SO<sub>2</sub> in atmosphere will be found and compared with experimental results obtained by means of direct laboratory measurement. Also there will be estimated the growth of SO<sub>2</sub> concentration caused by point emissions.

The paper is organized as follows: First, the method will be shortly described. Then, there will be presented results for time series from four measuring stations. Finally, we give an interpretation of obtained results and present the suggestions of further research.

## Method description

The method of the extraction of the system dynamics from stochastic time series has been presented in our previous paper [5]. Here, we limit introduction to short description of this method.

We assume, that the equation describing iterative evolution of the one-dimensional discrete system with noise is given by:

$$x_{n+1} = f(x_n) + g(x_n)\xi_n \quad (2)$$

To simplify the notation for analytical and numerical procedure of the method, we replace the eq. (2) by:

$$x' = f(x) + g(x)\xi \quad (3)$$

where  $f(x)$  is a deterministic part and  $g(x)$  is a stochastic part of the dynamics. The  $\xi$  denotes the noise term. The noise is uncorrelated with the mean equals to zero and standard deviation equals one, respectively. The functions  $f(x)$  and  $g(x)$  are determined from the formulas:

$$f(x) = \int_{-\infty}^{+\infty} x'q(x'|x)dx' \quad (4)$$

and

$$g(x) = \sqrt{\int_{-\infty}^{+\infty} x'^2q(x'|x)dx - f^2(x)} \quad (5)$$

where  $q(x'|x)$  is the conditional probability density, which is computed directly from time series by construction of histograms with the equal width bins size. The width of every bin is equal to the precision of measurement of concentration of the SO<sub>2</sub> (given in [ $\mu\text{g}/\text{m}^3$ ]),

which in our case is  $\Delta = 0.1$ . The procedure of obtaining of eq. (4) and (5) has been presented in details in [5]. Next, we search for all the pairs  $\{x_k, x_{k+1}\}$  for which  $x_k = x$  and  $x_{k+1} = x'$ . The number of these pairs is given by  $N_j$ . The  $N_i$  is the number of the points  $x_k = x$  in the time series. Thus, the conditional probability density is computed from:

$$q(x' | x) = \frac{N_j}{N_i \cdot \Delta} \quad (6)$$

Finally, the integrals in eq. (4) and eq. (5) may be computed numerically. In present paper a trapeze rule is employed.

## Data

Time series of  $\text{SO}_2$  concentrations recorded hourly are available in open-source database of Chief Inspectorate of Environmental Protection in Poland [18]. We focused on signals from four stations of 264 situated in Poland. For comparison of the processes of emission and of production of  $\text{SO}_2$  in different areas, we selected four stations, which codes are given below:

- MzGranicaKPN - located in Granica village, near the boundary of Kampinos National Park - the greatest forest complex in central Poland,
- PdBorsukowiz - located in Borsukowizna in Podlaskie region, which is known as "green lungs of Poland",
- MzWarWokalna - located in Warszawa (the biggest city in Poland) in Ursynow district, with high population density and high vehicular traffic,
- LdLodzCzerni - located in Lodz (the third biggest city in Poland), near its center, close to a few highways.

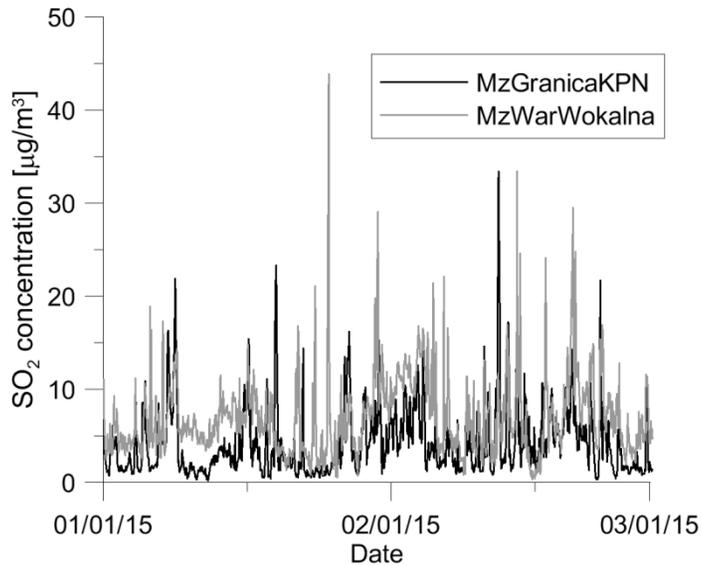


Fig. 1. Concentrations of  $\text{SO}_2$  from first two month of the year 2015 from measurement station MzGranicaKPN (background) and MzWarWokalna (urban)

Note, that two first stations are located in areas with relatively low population density and far from any industrial centers (further they are called “background stations”) while the next two are located near the centers of large cities (“urban stations”). Such selection and limitation of stations was also caused by three additional reasons: 1) for analysis we chose, the data with maximum 15 % missed measurements in year (which provides the adequate length of time series for the method) 2) the precision of measurement is 0.1, which was underlined in method description (requirement for unification of all datasets) 3) necessity to perform the analysis of concentration recordings from the same station from few years (for averaging and verification of repeatability of the results). As example we present concentrations of SO<sub>2</sub> from first two months of the year 2015 from MzGranicaKPN and MzWarWokalna in Figure 1.

## Results

Figure 2a,b presents function  $f$  and  $g$  from reconstruction of dynamics of SO<sub>2</sub> concentration in measurement station MzGranicaKPN (year 2015). While the  $g$  function fluctuates,  $f$  function may be approximated by a linear function with good accuracy. The similar linear trend is observed for all the considered data sets.

Figure 2c presents a histogram of noise terms, computed from eq. (2) as:

$$\xi_n = \frac{x_{n+1} - f(x_n)}{g(x_n)} \quad (7)$$

This distribution is slightly similar to a Gaussian one. However, some statistical tests we have done show that this is not normal distribution in fact. Problem of real type of noise distribution will be a subject of forthcoming investigations.

In Figure 3, we show the original time series of SO<sub>2</sub> concentration during first ten days of year 2015 and artificial series reconstructed from experimental SO<sub>2</sub> dynamics, basing on eq. (2) with noise terms obtained from (7). The agreement between these two series is quite satisfactory. The differences appear usually for extremal (minimal or maximal) values of concentration for which the accuracy of determination of functions  $f$  and  $g$  is the lowest.

Let us now analyze in details the deterministic function  $f$  extracted from time series.

By comparison (2) and (1), we express  $f$  function in form:

$$f(x) = (1 - k\Delta t)x + E\Delta t \quad (8)$$

where  $k$  and  $E$  should be both greater than zero, according to their interpretation given in Introduction. The form of eq. (8) justifies the linear form of function  $f$  which may be given in simply form  $f(x) = ax + b$  where  $a = 1 - k$ .

The slope  $a$  and intercept  $b$  have been determined by the least square method. We assess the quality of fit by determination coefficient  $R^2$ . In all performed fits, the  $R^2$  exceeds 0.96. To find the quality of fitting of linear function we have checked the autocorrelation of residuals by means of Durbin-Watson statistic [19]. For all the time series we have computed the value of  $d$  defined as:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (9)$$

where  $e_i$  denotes the residuum  $i$ -th bin and  $n$  is a number of bins. For all the time series analyzed we obtained the value of  $d$  between 1.1 and 1.4. It is known that there is no statistical evidence that the residuals are positively autocorrelated when  $d$  is above 1.2 for  $N = 20$  and 1.29 for  $N = 25$  (at significance level 0.05) [19]. According that we may conclude that in most cases the residuals have not been positively autocorrelated. However, there was some cases when test is not conclusive.

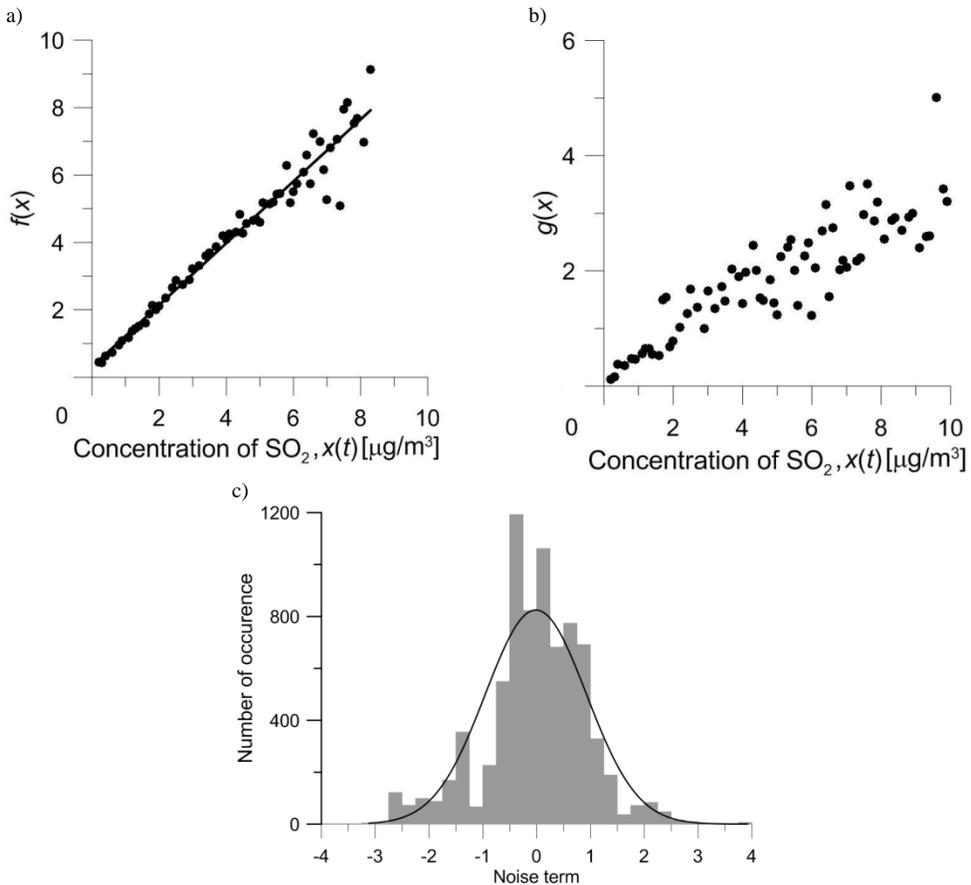


Fig. 2. Deterministic a)  $f(x)$  and stochastic b)  $g(x)$  part of reconstructed dynamics of  $\text{SO}_2$  concentration in MzGranicaKPN station (year 2015). c) Histogram of noise terms  $\xi$  obtained from reconstruction

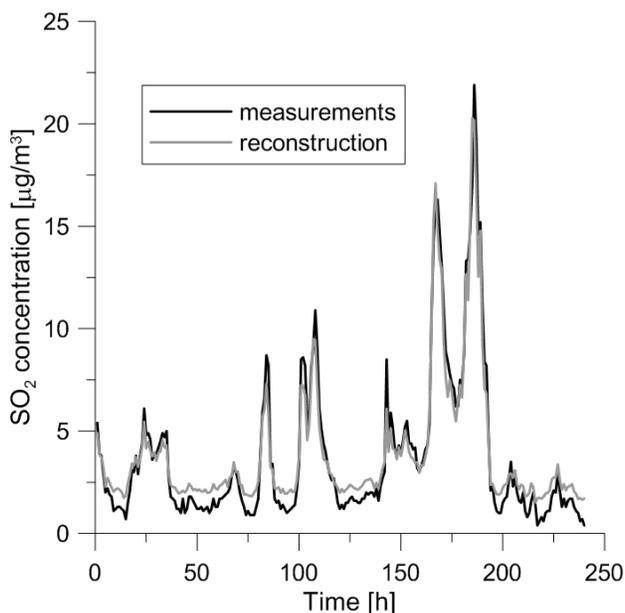


Fig. 3. Results of measurement and reconstruction of SO<sub>2</sub> concentration time series at the begin of year 2015 (MzGranicaKPN measurement station)

Note, that according to our previous observations [5], we limited the range of fit to points, in which number of counts ( $N_i$ ) in bin of histogram constructed for conditional probability density function was larger than 100 (see eq. (6)). The results are presented in Table 1.

Table 1  
Slope  $a$  and intercept  $b$  coefficients for deterministic part of eq. (2) defined as  $f(x) = ax + b$  for different measurement stations and different years

Measurement station	MzGranicaKPN		PdBorsukowiz		MzWarWokalna		LdLodzCzerni	
	slope $a$	intercept $b$						
2011	0.972	0.194	1.006	0.025	0.933	0.270	1.012	0.330
2012	0.961	0.304	0.937	0.113	0.930	0.541	0.975	0.481
2013	0.983	0.162	0.930	0.126	0.973	0.321	1.025	0.194
2014	0.998	0.136	0.919	0.141	0.914	0.544	0.975	0.423
2015	1.014	0.136	0.973	0.06	0.901	0.543	1.023	0.271
Mean	0.986	0.186	0.953	0.093	0.930	0.444	1.002	0.340
Standard deviation	0.021	0.070	0.036	0.049	0.027	0.137	0.025	0.115

We can recognize that for all the measuring series slope  $a$  is close to 1. For most of series (15 from 20) its value is slightly less than one. Also, the five-year mean values for three stations are less than one, thus we could estimate the decay rate for these stations. For one station (LdLodzCzerni) a value of slope is slightly greater than one, however the

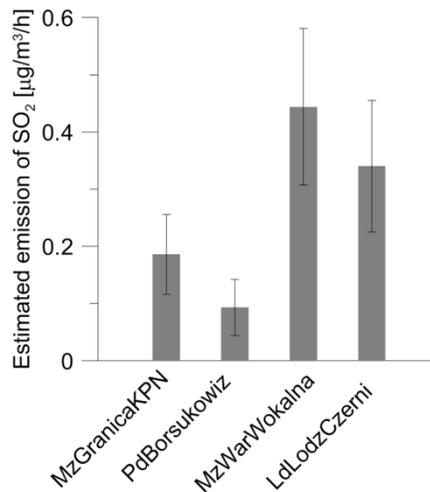
difference is much less than its standard deviations. Table 2 presents the values of decay rates and lifetimes for remained three stations.

Table 2

Decay rates and lifetimes of SO<sub>2</sub> for three measurement stations

Measurement station	MzGranicaKPN	PdBorsukowiz	MzWarWokalna
Decay rate $k$	$0.014 \pm 0.021 \text{ h}^{-1}$	$0.047 \pm 0.036 \text{ h}^{-1}$	$0.070 \pm 0.027 \text{ h}^{-1}$
Lifetime $\tau$	$71 \pm 107 \text{ h}$	$21 \pm 16 \text{ h}$	$14 \pm 5 \text{ h}$

Basing on these values we may estimate the mean lifetime as between 0 and 180 h that means less than a week. It is consistent with results of direct measuring of kinetics of SO<sub>2</sub> decay, according to which, SO<sub>2</sub> lifetime has been found in range from one day (scavenging of SO<sub>2</sub> by dry deposition) to one week (reaction with the OH radical at typical atmospheric levels of OH) [1]. Standard deviations in Table 2 indicate, that there is no distinct difference between the lifetime of SO<sub>2</sub> in urban and background conditions. The second coefficient of deterministic component may be interpreted as mean emission of SO<sub>2</sub>. There is a significant difference between two first and two last measurement stations. As we may see in Figure 4, emission of SO<sub>2</sub> in urban stations is estimated approximately two and a half times greater than in background stations.

Fig. 4. Estimated emission of SO<sub>2</sub> for four measuring stations

To confirm the differences between the background and urban environment we have used Welch's  $t$ -test, which is suitable to test the hypothesis that two data sets have equal means [20]. According to that test we have computed the statistic  $t$  by means of formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (10)$$

where  $\bar{X}_i$ ,  $s_i$  and  $n_i$  denotes the  $i$ -th sample mean, population variance and sample size, respectively (the last one was equal to 5 for all the samples while we analyzed concentration series from five successive years). According to Welch's  $t$ -test procedure, this value has at-Student distribution with number of degrees of freedom equal to:

$$v = \frac{\left( \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}} \quad (11)$$

This result is in agreement with expectations, according to which, emission of SO<sub>2</sub> in great industrial centers should be greater than that observed outside of these centers. By use of this test, we have found that at significance level 0.05 mean values of estimated emission of SO<sub>2</sub> between background and urban stations differ. This difference has been found for every pair of background and urban station. Moreover, the mean values between two background stations (MzGranicaKPN and PdBorsukowiz) also differ. However, the hypothesis that the mean values for two urban stations are equal is not rejected at the same significance level.

## Conclusions

In this paper, we applied the stochastic approach for experimental time series of SO<sub>2</sub> concentration. This method was used for reconstruction of the time series to separate the deterministic and stochastic factors influencing the dynamics of SO<sub>2</sub> measurement series. The dynamics of SO<sub>2</sub> concentration after removal the noise term (which may arise from such random factors as fluctuation of wind or air moisture) has been described as a sum of positive mean source term and first order decay. In many studies [1] it has been shown that SO<sub>2</sub> in troposphere undergoes many reactions which usually have form of first order decay. The lifetime of SO<sub>2</sub> has been estimated for few days, which is also consistent with experimental results on SO<sub>2</sub> reactions in troposphere. Moreover, it has been found that there is not significant difference between lifetime of SO<sub>2</sub> in huge industrial regions and outside of them.

Our analysis have shown significant difference between emission levels measured in stations located in big cities and outside of them. The emission near urban stations has been estimated as a few time greater than that in the vicinity of background ones.

We conclude, that the proposed method is not a universal tool for investigations of dynamics of pollutants in atmosphere. However, the interpretation of dynamics components can be directly related to mechanisms of reactions, which dominate in observed changes of concentrations. The length of the data is the method limitation. In our previous works, we have estimated the minimum length of time series for 10000 points. It means, that e.g. for concentration measured in every hour the minimal time period for analysis is one year. In result, we are not able to study the seasonal changes of the dynamics of SO<sub>2</sub> concentration. From stochastic analysis, we obtain the dynamics of the system averaged over the whole year.

However the method of stochastic dynamics reconstruction is not limited to analysis of SO<sub>2</sub> concentration. It is a promising tool, which can be applied for other ingredients of troposphere. The investigations in this subject are now in progress.

## Acknowledgement

This work has been partially financed by the project of the National Science Centre (contract No. 2016/21/B/ST10/02998).

## References

- [1] Seinfeld JH, Pandis SN. Atmospheric Chemistry and Physics. From Air Pollution to Climate Change. Hoboken: John Wiley & Sons; 2016. ISBN: 9781118947401.
- [2] Stern AC, editor. Fundamentals of Air Pollution. New York: Academic Press; 2014. ISBN: 9780123736154.
- [3] Smith KR, Bruce N, Balakrishnan K, Adair-Rohani H, Balmes J, Chafe Z, et al. Millions dead: How do we know and what does it mean? methods used in the comparative risk assessment of household air pollution. *Ann Rev Public Health*. 2014;35:185-206. DOI: 10.1146/annurev-publhealth-032013-182356.
- [4] Sunyer J, Urban air pollution and chronic obstructive pulmonary disease: a review. *European Respiratory J*. 2001;17(5):1024-1033. <https://erj.ersjournals.com/content/17/5/1024.long>.
- [5] Petelczyc M, Żebrowski JJ, Gac JM. Extraction of stochastic dynamics from time series. *Phys Rev E*. 2012;86:011114. DOI: 10.1103/PhysRevE.86.011114.
- [6] Olszowski T. Seasonal values of the gaseous concentrations of air quality ratings in a background area. *Ecol Chem Eng S*. 2013;20(4):719-732. DOI: 10.2478/eces-2013-0050.
- [7] Hosseinibalam F, Ghaffaripasand O. The effects of emission sources and meteorological factors on sulphur dioxide concentration of Great Isfahan, Iran. *Atmos Environ*. 2015;100:94-101. DOI: 10.1016/j.atmosenv.2014.10.012.
- [8] Meng Z-Y, Xu X-B, Wang T, Zhang X-Y, Yu X-L, Wang S-F, et al. Ambient sulfur dioxide, nitrogen dioxide, and ammonia at ten background and background sites in China during 2007-2008. *Atmos Environ*. 2010;44:2625-2631. DOI: 10.1016/j.atmosenv.2010.04.008.
- [9] Lee C, Martin RV, van Donkelaar A, Lee H, Dickerson RR, Hains JC, et al. SO<sub>2</sub> emissions and lifetimes: Estimates from inverse modeling using in situ and global, space-based (SCIAMACHY and OMI) observations. *J Geophys Res*. 2011;116:D06304. DOI: 10.1029/2010JD014758.
- [10] Khoder MI. Atmospheric conversion of sulfur dioxide to particulate sulfate and nitrogen dioxide to particulate nitrate and gaseous nitric acid in an urban area. *Chemosphere*. 2002;49:675-684. DOI: 10.1016/S0045-6535(02)00391-0.
- [11] Herrmann H, Ervens B, Jacobi H-W, Wolke R, Nowacki P, Zellner R. CAPRAM2.3: A chemical aqueous phase radical mechanism for tropospheric chemistry. *J Atmos Chem*. 2000;36:231-284. DOI: 10.1023/A:1006318622743.
- [12] Hoigné J, Bader H, Haag WR, Staehelin J. Rate constants of reactions of ozone with organic and inorganic compounds in water. III - Inorganic compounds and radicals. *Water Res*. 1985;19:993-1004. DOI: 10.1016/0043-1354(85)90368-9.
- [13] McArdle JV, Hoffmann MR. Kinetics and mechanism of the oxidation of aquated sulfur dioxide by hydrogen peroxide at low pH. *J Phys Chem*. 1983;87:5425-5429. DOI: 10.1021/j150644a024.
- [14] Lind JA, Lazrus AL. Aqueous-phase oxidation of sulfur(IV) by some organic peroxides. *Eos Trans*. 1983;64:670-675. DOI: 10.1029/JD092iD04p04171.
- [15] Lind JA, Lazrus AL, Kok GL. Aqueous phase oxidation of sulfur(IV) by hydrogen peroxide, methyl hydroperoxide and peroxyacetic acid. *J Geophys Res*. 1987;92:4171-4177. DOI: 10.1029/JD092iD04p04171.
- [16] Misra CD, Khanooja PS, Sharma AK, Mudgal PK, Bansal SP, Gupta KS. Studies on rainwater and dry depositions at Jaipur, India. *J Ind Chem Soc*. 2013;90:1137-1146. [https://www.researchgate.net/publication/287410793\\_Studies\\_on\\_rainwater\\_and\\_dry\\_depositions\\_at\\_Jaipur\\_India](https://www.researchgate.net/publication/287410793_Studies_on_rainwater_and_dry_depositions_at_Jaipur_India).
- [17] Zhang L, Brook JR, Vet R. A revised parameterization for gaseous dry deposition in air-quality models. *Atmos Chem Phys*. 2003;3:2067-2082. DOI: 10.5194/acp-3-2067-2003.
- [18] <http://powietrze.gios.gov.pl/pjp/archives>.
- [19] Durbin J, Watson GS. Testing for serial correlation in least squares regression, I. *Biometrika* 1950;37:409-428. DOI: 10.1093/biomet/37.3-4.409.
- [20] Welch BL. The generalization of "Student's" problem when several different population variances are involved. *Biometrika*. 1947;34:28-35. DOI: 10.1093/biomet/34.1-2.28.