

# Monitoring and Controlling the Air Quality- Environmental Pollutants. Evaluation of Uncertainty Budget

Fănel Iacobescu<sup>1</sup>, Alexandru Sălceanu<sup>2</sup>, Albert-Ciprian Anghel<sup>3</sup>,  
Mirela-Adelaida Anghel<sup>1</sup>

<sup>1</sup> *Romanian Bureau of Legal Metrology, Vitan-Bârzești, 11, office@brml.ro,  
mirela.a.anghel@gmail.com,*

<sup>2</sup> *Gheorghe Asachi Technical University of Iași, Dimitrie Mangeron, e-mail address(es)*

<sup>3</sup> *University of Bucharest, 4-12 Queen Elisabeth Avenue, sect. 3, albert.ciprian@gmail.com*

**Abstract** – Accurate and reliable air pollutant measurements are widely required all over the world. Measurement science is something of vital importance to all of us. The monitored pollutants, the measurement methods, the limit values, the alert thresholds and the criteria for locating and monitoring points have been established by national legislation on the protection of the atmosphere and meet the requirements stipulated by EU regulations. The paper presents the conclusions which have been drawn using statistical inference tools, based on the data collected during measurements of few air pollutants CO, SO<sub>2</sub>, NO<sub>x</sub>. Measurements were collected for a period of 3 months, 10 days / month, using gas analysers configured with NDIR (Non – Dispersive Infrared Absorbtion) sensors. Additionally, based on experience, the paper describes, step by step, an example of uncertainty budget evaluation of SO<sub>2</sub> for measuring air pollutants.

**Key words:** environment protection, metrology, air pollutants, uncertainty evaluation

## I. INTRODUCTION

Quality of environmental air is directly linked to the quality of life. It is compulsory for everybody to respect de humanity and the nature; too many governmental and non-governmental bodies, high tech equipment manufacturers, independently laboratories are involved in standardisation, law enforcement, monitoring and measuring the quality of air. In accordance with Law no. 265/2006 approving the Governmental Emergency Ordinance 195/2005 on environmental protection responsibility for monitoring ambient air quality in Romania belong to environmental authorities.

## II. MONITORING SYSTEM

In Romania, 142 air quality monitoring stations have been equipped with automatic instruments which measure the main important atmospheric pollutants.

All these measured data are sent to 41 Environmental Protection Agencies, and then, these data are displayed by means of exterior panels, conventionally located within the densely populated areas of the cities and by means of interior panels, located in city halls.



*Fig. 1. Traffic type*



*Fig. 2. Industrial type*



*Fig. 3. Urban background type*



*Fig. 4. Suburban background type*



Fig. 5. Regional background type



Fig. 6. EMEP type



Fig. 7. Exterior panel



Fig. 8. Interior panel

Central database collects and archives the primary data, validated data and certified data, as well.

This database is accessed by specialists in order to make different studies and to send our national periodical reports to European level.

For example, the limit values of sulphur dioxide (SO<sub>2</sub>) air pollutant, according to law no. 104 from June 15<sup>th</sup> 2011, are: **500 µg/m<sup>3</sup> - the limit value for three hours** of continuous monitoring of air quality, in the representative points of an area of at least 100 km<sup>2</sup>, an entire area or a crowded area; **350 µg/m<sup>3</sup> - hourly limit value** for human health protection; **125 µg/m<sup>3</sup> - daily limit value** for human health protection; **20 µg/m<sup>3</sup> - critical level** (annual limit value) for vegetation protection [5].

### III. BASIC CONCEPT ON INFERENCE AND STATISTICAL TOOLS USED IN DATA PROCESSING

This chapter presents the statistical tools used to conduct hypothesis tests on the regression coefficients obtained in linear regression and the significance of regression for the measured data [5]. Statistical inference is based on the laws of probability, and allows analysts to infer conclusions about a given population based on results observed or measured [6]. Due to the fact that statistical methods have proven their importance for researchers interested to interpret the results of their experiments it was the selected way to design those experiments which are able to provide interpretable results.

In general, a linear regression model is described by the following equation (1):

$$Y = X\mathbf{m} + \varepsilon \quad (1)$$

On the components of vector error is the assumption that dispersion is an exponential function of  $p \in \mathbf{N}^*$ ,  $p \leq k$ , control variables of the form:

$$D(\varepsilon_i) = M(\varepsilon_i^2) = \sigma_i^2 = \exp[\mathbf{z}_i^T \cdot \boldsymbol{\alpha}], \quad i = \overline{1, n} \quad (2)$$

where:

$\mathbf{z}_i^T = (z_{i1}, z_{i2}, \dots, z_{ip})$  is a known line vector and  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$  is an unknown column vector.

Presuming that  $z_{i1} = 1 \quad \forall i = \overline{1, n}$  then dispersion errors will have the following form:

$$D(\varepsilon_i) = \exp(\alpha_1) \cdot \exp(\alpha_2 z_{i2}) \cdot \dots \cdot \exp(\alpha_p z_{ip}), \quad i = \overline{1, n} \quad (3)$$

which gives the reason why this model is called the multiplicative heteroscedasticity.

There are some tests which allow verifying the heteroscedasticity hypothesis. The Goldfeld – Quandt test check the  $H_0$  hypothesis against the alternative one,

$H_1: H_1: \sigma_1^2 \leq \sigma_2^2 \leq \dots \leq \sigma_n^2$  through the following steps:

- 1) Assuming that  $H_1$  is true, the selection values are ordered in ascending order of dispersion errors;
- 2)  $r$  central values of the selection are deleted;
- Obs. Sometimes it is useful to work with  $r = 0$ . (W.Green)
- 3) the remaining selection values are divided into two groups, one formed by the first  $\frac{1}{2}(n-r)$  values and other

with the latest  $\frac{1}{2}(n-r)$  selection values. It is necessary to

fulfil the following conditions:  $\frac{1}{2}(n-r) \in \mathbf{N}^*$  and

$\frac{1}{2}(n-r) \geq k$ , where  $k$  is the number of regressors of the model.

- 4) For each group is calculated the sum of squares error, and then is calculated the statistic  $\lambda = S_1 / S_2$ , where  $S_1$  is

the higher of two amounts, and  $S_2$  is the other.

Obs. If the null hypothesis is true, it can be shown that  $\lambda$  has a distribution  $F$  with  $\frac{1}{2}(n-r) - (k+1)$  and

$\frac{1}{2}(n-r) - (k+1)$  degrees of freedom.

- 5) For a given significance level  $\beta$  the critical value  $F_c$  is determined from the  $F$  distribution table and compared with  $\lambda$ .

If  $\lambda < F_c$  the null hypothesis is accepted and the  $H_1$

hypothesis is rejected with an error given by  $\beta$ .

Otherwise, we accept the heteroscedasticity hypothesis and reject the null hypothesis with an error given by  $\beta$ .

The table 1 presents the statistical coefficient  $\lambda$  calculated for each group of measurement results, obtained in 24 hours by each monitoring station, part of Romanian ambient air quality monitor network.

### IV. UNCERTAINTY BUDGET EVALUATION

Confidence is vital in enabling metrology to link human activities together across geographic and professional boundaries. This confidence becomes enhanced with the increased use of network co-operation,

Table 1. Statistical coefficient  $\lambda$  calculated for measurements results over 24 hours by each monitor station

	$S_{AB-1}^2$	$S_{AB-2}^2$	$S_{AG-1}^2$	$S_{AG-2}^2$	$S_{AG-6}^2$	$S_{AR-1}^2$	$S_{AR-2}^2$	$S_{BC-1}^2$	$S_{BC-2}^2$	$S_{BC-3}^2$	$S_{BT-1}^2$	$S_{CI-4}^2$	$S_{CI-5}^2$	$S_{CS-3}^2$	$S_{CS-4}^2$	$S_{CT-3}^2$	$S_{CT-4}^2$	$S_{CV-1}^2$	$S_{DI-1}^2$
$S_{AB-1}^2$	1	1.28	87.73	41.97	35.39	59.97	45.48	53.38	82.63	62.41	107.13	12.01	8.66	82.89	148.58	76.96	553.36	3.33	40.79
$S_{AB-2}^2$	0.78	1	68.53	32.79	27.64	46.85	35.52	41.70	64.54	48.75	83.69	9.38	6.76	64.75	116.06	60.12	432.26	2.60	31.86
$S_{AG-1}^2$	0.01	0.01	1	0.48	0.40	0.68	0.52	0.61	0.94	0.71	1.22	0.14	0.10	0.94	1.69	0.88	6.31	0.04	0.46
$S_{AG-2}^2$	0.02	0.03	2.09	1	0.84	1.43	1.08	1.27	1.97	1.49	2.55	0.29	0.21	1.97	3.54	1.83	13.18	0.08	0.97
$S_{AG-6}^2$	0.03	0.04	2.48	1.19	1	1.69	1.29	1.51	2.34	1.76	3.03	0.34	0.24	2.34	4.20	2.18	15.64	0.09	1.15
$S_{AR-1}^2$	0.02	0.02	1.46	0.70	0.59	1	0.76	0.89	1.38	1.04	1.79	0.20	0.14	1.38	2.48	1.28	9.23	0.06	0.68
$S_{AR-2}^2$	0.02	0.03	1.93	0.92	0.78	1.32	1	1.17	1.82	1.37	2.36	0.26	0.19	1.82	3.27	1.69	12.17	0.07	0.90
$S_{BC-1}^2$	0.02	0.02	1.64	0.79	0.66	1.12	0.85	1	1.55	1.17	2.01	0.22	0.16	1.55	2.78	1.44	10.37	0.06	0.76
$S_{BC-2}^2$	0.01	0.02	1.06	0.51	0.43	0.73	0.55	0.65	1	0.76	1.30	0.15	0.10	1.00	1.80	0.93	6.70	0.04	0.49
$S_{BC-3}^2$	0.02	0.02	1.41	0.67	0.57	0.96	0.73	0.86	1.32	1	1.72	0.19	0.14	1.33	2.38	1.23	8.87	0.05	0.65
$S_{BT-1}^2$	0.01	0.01	0.82	0.39	0.33	0.56	0.42	0.50	0.77	0.58	1	0.11	0.08	0.77	1.39	0.72	5.17	0.03	0.38
$S_{CI-4}^2$	0.08	0.11	7.30	3.49	2.95	4.99	40.51	4.44	6.88	5.20	8.92	1	0.72	6.90	12.37	6.41	46.07	0.28	3.40
$S_{CI-5}^2$	0.12	0.15	10.13	4.85	4.09	6.93	68.65	6.17	9.54	7.21	12.37	1.39	1	9.57	17.16	8.89	63.92	0.38	4.71
$S_{CS-3}^2$	0.01	0.02	1.06	0.51	0.43	0.72	52.06	0.64	1.00	0.75	1.29	0.14	0.10	1	1.79	0.93	6.68	0.04	0.49
$S_{CS-4}^2$	0.01	0.01	0.59	0.28	0.24	0.40	61.11	0.36	0.56	0.42	0.72	0.08	0.06	0.56	1	0.52	3.72	0.02	0.27
$S_{CT-3}^2$	0.01	0.02	1.14	0.55	0.46	0.78	94.58	0.69	1.07	0.81	1.39	0.16	0.11	1.08	1.93	1.00	7.19	0.04	0.53
$S_{CT-4}^2$	0.00	0.00	1.16	0.08	0.08	0.11	71.44	0.10	0.15	0.11	0.19	0.02	0.02	0.15	0.27	0.14	1	0.01	0.07
$S_{CV-1}^2$	0.30	0.38	26.36	12.61	10.63	18.02	122.63	16.04	24.83	18.76	32.19	3.61	2.60	24.91	44.65	23.13	166.29	1	12.26
$S_{DI-1}^2$	0.02	0.03	2.15	1.03	0.87	1.47	13.75	1.31	2.03	1.53	2.63	0.29	0.21	2.03	3.64	1.89	13.57	0.08	1

common units of measurement and common measuring procedures, as well as the recognition, accreditation and mutual testing of measuring standards and laboratories in different countries. Mankind has thousands of years of experience confirming that life really does become easier when people co-operate on metrology.

Especially when we talk about pollutants in the environment we have to take into consideration traceability and interlaboratory comparison of measurement performed on the air, which cross freely any border.

The following example refers to a gas analyser designed and manufactured to identify qualitative and determine quantitative the sulphur dioxide ( $SO_2$ ) concentrations from the ambient air. The analyser has the following declared technical specifications:

- Measurement range: (0...500) ppb;
- Resolution: 1 ppb;
- Noise: 0.1 % of the measured concentration;
- Zero drift due to temperature variation 0.1 ppb / °C;
- Zero drift after 30 days: 0.5 % of the measured concentration;
- Span drift due to temperature variation: 0.1 ppb / % / °C;
- Span drift after 30 days: 0.5 % of the measurement concentration;
- Linearity error:  $\pm 1$  % of the full scale;
- Reproducibility: 1 % of the measured concentration.

#### A. Data processing

For the purpose of this experiment, the calibration of considered  $SO_2$  gas analyser, was performed in one point, using a standard gas mixture, having a concentration of

(105  $\pm$  2) ppb  $SO_2$  / air, at the end of 30 days, under continuous monitoring. The temperature in the chamber was maintained stabilized at 20 °C.

The volumetric fractions displayed by the  $SO_2$  gas analysers, expressed in ppb, are the following:

102 103 104 105 102 102 104 105 103 105

**Note 1:** The conversion of different measurements units, in standard environmental conditions (1.013 bar; 20 °C), used for  $SO_2$  is given below: 1 ppb = 0.001 ppm =  $1 \times 10^{-7}$  % (volumetric fraction) =  $2.666 \cdot 10^{-3}$  mg / m<sup>3</sup> = 2.666 3  $\mu$ g / m<sup>3</sup>.

#### B. Uncertainty budget for $SO_2$ gas analyser

The uncertainty budget, calculated for  $SO_2$  analyser according to the following formula is presented in table 2.

$$\begin{aligned}
 \gamma_{SO_2} = & \overline{\gamma}_{SO_2} + f_{CE} + f_{noise} + f_{resolution} + f_{\text{drift}} + \\
 & f_{3\text{Cdrift}} + f_{\text{err}} + f_{\text{linearity}} + \\
 & + f_{\text{Reproductibility}} + f_{3\sigma_{\text{err}}}
 \end{aligned} \tag{4}$$

**The measurement result** is reported as:  $y = \bar{y} \pm U$ , after replacing the numerical values, it becomes: **(103.5  $\pm$  6.4) ppb,  $k=2$**

**Note 2:** Generally, the uncertainty value will round up, with the exception of those uncertainty values smaller than 5 %, which will be rounded down.

**Note 3:** As a rule, the measurement uncertainty is expressed by 1 or 2 significant figure(s). The result of measurement will be round so that its last digit(s) has the same order as the uncertainty figures.

In order to draw conclusions related to the data collected during measurements of few air pollutants CO, SO<sub>2</sub>, NO<sub>x</sub> it was necessary to use statistical inference tools. Measurements have been collected for a period of 3 months, 7 days / month, 24 hours / day, using gas analysers configured with NDIR sensors, sensors based on NonDispersive Infrared Absorption principle. For each group of measurements (CO, SO<sub>2</sub>, NO<sub>x</sub>) have been calculated separately the average and the sum of

squares error. In order to check  $H_0$  hypothesis against the alternative one  $H_1 : \sigma_1^2 \leq \sigma_2^2 \leq \dots \leq \sigma_n^2$  has been applied the Goldfeld–Quandt test [6], [7].

For a given significance level of 95 % is determined critical value  $F_c$  from the  $F$  distribution table and compared with  $\lambda$  calculated for each sets of data generated by CO, SO<sub>2</sub>, NO<sub>x</sub> measurements.

The red values represent those values for which the heteroscedasticity hypothesis is rejected.

Table 2. Uncertainty budget calculated for SO<sub>2</sub> gas analyser

Quantity, ppb	Value	Uncertainty due to	Probability distribution	Standard uncertainty
$\overline{y}_{SO_2}$	103.4	<ul style="list-style-type: none"> <li>measurement repetability</li> </ul> $u_{repeatability} = \frac{s_{SO_2}}{\sqrt{n}} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n(n-1)}}$	Normal	0.401
$f_{CE}$	1	<ul style="list-style-type: none"> <li>standard gas mixture from the calibration certificate, where <math>U(SO_2)_{CE} = 2</math> ppb (k=2),</li> </ul> $u_{CE} = \frac{U_{CE}}{k}$	Normal	1.000
$f_{noise}$	1	<ul style="list-style-type: none"> <li>due to noise</li> </ul> $u_{noise} = \frac{\Delta noise}{2\sqrt{3}} = \frac{0.1 \cdot 105}{100 \cdot 2\sqrt{3}}$	Rectangular	0.030
$f_{resolution}$	1	<ul style="list-style-type: none"> <li>limited rezolution</li> </ul> $u_{resolution} = \frac{\Delta resolution_{SO_2}}{2\sqrt{3}} = \frac{1}{2\sqrt{3}}$	Rectangular	0.289
$f_{0\_drift}$	1	<ul style="list-style-type: none"> <li>temperature variation – zero drift</li> </ul> $u_{0\_drift} = \frac{\alpha \cdot \Delta T}{\sqrt{2}} = \frac{0.1 \cdot 2}{\sqrt{2}}$	Arcsin	0.141
$f_{30\_drift}$	1	<ul style="list-style-type: none"> <li>temperature variation – span drift (after 30 days)</li> </ul> $u_{30\_drift} = \frac{\alpha \cdot x_{CE} \cdot \Delta T}{\sqrt{2}} = \frac{0.1 \cdot 105 \cdot 2}{100 \cdot \sqrt{2}}$	Arcsin	0.148
$f_{liniarity}$	1	<ul style="list-style-type: none"> <li>liniarity of the device</li> </ul> $u_{liniarity} = \frac{\varepsilon_l(\%)}{100} \cdot x_{max\_range} \cdot \frac{1}{\sqrt{3}} = \frac{1}{100} \cdot 500 \cdot \frac{1}{\sqrt{3}}$	Rectangular	2.887
$f_{reproductibility}$	1	<ul style="list-style-type: none"> <li>reproductibility</li> </ul> $u_{reproductibility} = \frac{\varepsilon_R(\%)}{100} \cdot x_{CE} \cdot \frac{1}{\sqrt{6}} = \frac{1}{100} \cdot 105 \cdot \frac{1}{\sqrt{6}}$	Trianghiular	0.429
$f_{0\_err}$	1	<ul style="list-style-type: none"> <li>zero error</li> </ul> $u_{0\_err} = \frac{1}{\sqrt{3}}$	Rectangular	0.577
$f_{30\_err}$	1	<ul style="list-style-type: none"> <li>span error (after 30 days)</li> </ul> $u_{30\_err} = \frac{1}{\sqrt{3}}$	Rectangular	0.577
$\overline{y}_{SO_2}$	103.4	<p>The combined uncertainty was calculated [2] taking into account all input parameters presented above and is given by:</p> $u_{c\_SO_2} = \sqrt{u_{repeatability}^2 + u_{CE}^2 + u_{noise}^2 + u_{resolution}^2 + u_{0\_drift}^2 + u_{30\_drift}^2 + u_{liniarity}^2 + u_{reproductibility}^2 + u_{0\_err}^2 + u_{30\_err}^2}$	Normal	3.236

The conclusion of this heteroscedasticity test is that, for the significance level of 95 %, the measurements belong to the same population. For two days, all 142 stations for monitoring of air quality had measured lower concentrations of pollutants than in normal days, but the measurements have been comparable all over the country. The nature is unpredictable and is strongly depending on the pollution sources. Based on the database of measurements, on associated uncertainty and on statistical inference method can be drawn the conclusion that the monitoring stations are offering reliable measurements.

## V. CONCLUSIONS

Many physical and chemical toxic substances have an immediate impact on the quality of the world in which we live, so it is important to perform accurate measurements of their concentration. We measure pollutant gases in order to obtain data which can then be used to help make the correct decisions to properly care of our environment. The quality of our life depends on reliable measurements. Data quality and testing are therefore an important part of many regulations. There are many areas of legislation, where measurements are required to assess conformity with regulations or legislation e.g. healthcare, environmental and pollution control and so on. Uncertainty is a quantitative measure of the quality of measurement result, enabling the measurement results to be compared with other results, references, specifications or standards. The uncertainty increases down the traceability chain.

Currently, our national monitoring network is consisting of 142 stations for monitoring of air quality, equipped with automatic measurement of main pollutants from ambient air as follows: SO<sub>2</sub>, NO<sub>x</sub> (NO / NO<sub>2</sub>), CO, O<sub>3</sub>, PM10, PM 2.5, Pb, Cd, As, PAHs. The measurements registered by all 142 monitoring stations over a period of 3 months, 7 days / month, 24 hours / day have been reported with associated uncertainty. The conclusions which have been drawn using statistical inference tools are that monitoring stations are part of a coherent measurement system which is able to offer reliable and traceable results.

## VI. REFERENCES

- [1] International Vocabulary of Metrology, "Basic and General Concepts and Associated Terms" VIM, 3<sup>rd</sup> edition, JCGM 200:2008.
- [2] "Evaluation of measurement data – Guide to the expression of uncertainty in measurement", JCGM 100:2008.
- [3] <http://www.bipm.org>.
- [4] Metrology – in short, 3<sup>rd</sup> edition, July 2008, Euramet.
- [5] <http://www.calitate aer.ro>.
- [6] M.Evans, N.Hastings, B.Peacock, "Statistical distributions", Wiley, 2000.
- [7] J.R.Rice, "Mathematical Statistics and Data Analysis", second ed. Duxbury Press, Belmont, Ca., USA, 1995.