International Journal of Mechanical Engineering and Applications
2023; 11(3): 66-73
http://www.sciencepublishinggroup.com/j/ijmea
doi: 10.11648/j.ijmea.20231103.12
ISSN: 2330-023X (Print); ISSN: 2330-0248 (Online)

Research on Reliability of Online Monitoring Results of Factory Discharge Pollutant SO_2 Based on Regression Analysis

Yang Shuo[1], Pan Zhiqiang[2], Niu Xingrong[3], Geng Lei[3], Sun Zhijing[4], Ma Dongchang[5], Wang Douwen[6, *]

[1]Shahekou District Government Office, Dalian, China

[2]Operations Division of Westpac Petrochemical Company, Dalian, China

[3]China National Accreditation Service for Conformity Assessment, Beijing, China

[4]Dalian Product Quality Inspection and Testing Institute Co., Ltd., Dalian, China

[5]Zhongren Guozhi (Beijing) Testing Technology Research Institute, Beijing, China

[6]Customs Technology Center, Dalian, China

Email address:

4595771@qq.com (Yang Shuo), panzhiqiang@petrochina.com.cn (Pan Zhiqiang), niuxr@cnas.org.cn (Niu Xingrong), gengl@cnas.org.cn (Geng Lei), 531639538@qq.com (Sun Zhijing), zrgzyjy@126.com (Ma Dongchang), wangdouwen@163.com (Wang Douwen)

[*]Corresponding author

To cite this article:

Yang Shuo, Pan Zhiqiang, Niu Xingrong, Geng Lei, Sun Zhijing, Ma Dongchang, Wang Douwen. Research on Reliability of Online Monitoring Results of Factory Discharge Pollutant SO_2 Based on Regression Analysis. *International Journal of Mechanical Engineering and Applications*. Vol. 11, No. 3, 2023, pp. 66-73. doi: 10.11648/j.ijmea.20231103.12

Received: June 18, 2023; **Accepted**: July 10, 2023; **Published**: August 4, 2023

Abstract: Concerning the SO_2 pollution source monitoring, discussed and even debated all the time, is in the environmental field. Adhere to purpose of "lucid waters and lush mountains are invaluable assets", the immediate task is to establish a complete and correct QA/QC monitoring system. In China, there are a large number of online devices, for its superiority compared with the laboratory technology, that undertake tests. However, it also has to be admitted that, the online system, belonging to a non-standard, shall paid more attention to its effectiveness. In this paper, a Deming regression technique of variable error model, with unbiased correction (CSS_0), constant bias correction (CSS_1) and linear bias correction (CSS_2) step by step, is used to fit at levels between online and its standard system. F and t, as well as $\chi 2$ distribution test are subsequently followed by for the selected CSS. Finally, under the independent identical distribution (i.i.d) condition based on the bias correction, use A^* test to predict series residuals, from the correction, for its i.i.d condition. The uncertainty assessment, brought by the correction under site precision, combines the various variation to the maximum extent, and avoid the complicated correlation, is helpful for the quality assurance of the online system.

Keywords: Deming's Regression, Closeness Sum of Squares, Weighted Fitting, Residuals, $s_{R'}$, A^* Test

1. Introduction

Although the author is engaged in online field management decision-making work, but prefer to think on the technical level. Concerning the problem of metrological traceability of air pollution source monitoring in the environmental field, the author has discussed and even debated with Chinese professionals for many times in recent days. However, the author insists that economic development and environmental protection should be placed equal emphasis on the concept of "lucid waters and lush mountains are invaluable assets". The urgent task now is to establish a complete and correct air quality QA/QC monitoring system as soon as possible.

In our country, petroleum, chemical, steel and other fields have a considerable number of sets of online monitoring equipment. Different from the single analysis of indoor sampling and single configuration, the online analysis of automatic monitoring has entered the field under harsh environmental conditions from the indoor, and then enters the process from the site. This kind of analysis technology can quickly adapt to the multi-component environmental protection field will soon become the mainstream. The author acknowledges that this analysis system can collect the working parameters and intermediate information (pressure, temperature, flow, etc.) of the equipment in the process, and can effectively run with full load and provide high quality data in line with the user. Compared with the expensive physical and chemical analysis indoors, the process control of this system brings great technical benefits to the field of environmental protection detection.

However, the author also has to admit, this kind of online analysis basically belongs to a non-standard system of detection technology, with the lag of our online analysis trend development, more attention should be paid to the effectiveness of this non-standard system confirmation and uncertainty study.

In this paper, we propose a Deming regression technique [1] of variable error model, which is different from OLR (ordinary least squares regression) method used in our country. It also emphasizes that the determination of minimum bias without practical significance is not a statistical decision, but a subjective decision directly dependent on the user's application requirements. In other words, as long as the weighted orthogonal regression technical route given in this study is followed, there is no need to worry about the statistically meaningless bias correction amount, and the direction and objectives of this research can be successfully completed.

2. Overview

Clean air is one of the environmental factors that human beings depend on for survival. With the industrial, especially the large use of coal and oil, will produce a large number of harmful substances and soot, sulfur dioxide (SO_2), nitrogen oxides (NOx), carbon monoxide (CO), hydrocarbons and other emissions into the atmosphere, when the concentration exceeds the limit allowed by the environment and lasts for a certain time, it will change the normal composition of the air, destroy the natural ecological balance system, Thus endangering people's life, work and health, damaging natural resources and property. Or air pollution.

This kind of unorganized emission and unit perimeter gas molecular state primary pollution source diffusion has the characteristics of time distribution, space distribution, pollution factor intensity and comprehensive effect. The concentration of air pollution caused by the same pollution source to the same place at different times is often several times to tens of times different; Different places at the same time vary greatly. Due to the limitation of inversion layer, air temperature and air pressure, the concentration of primary pollutants is higher in the morning and evening, but lower at noon. If the wind speed is large and the atmosphere is unstable, the dilution and diffusion rate of pollutants are fast and the concentration changes quickly. On the contrary, dilution diffusion is slow and concentration changes slowly.

In view of the temporal and spatial characteristics of environmental pollution, it is only possible to reveal its change rule and predict its change trend and accuracy from large sample data by relying on the online continuous and tracking monitoring with superior technology and management, and entrusting the correct monitoring system tracking (representative selection of monitoring network and point).

This paper focuses on the routine or routine monitoring of SO_2 emission status of plant sources, evaluating the effect of control measures, measuring the implementation of environmental standards and the progress of environmental protection work. In order to study environmental capacity, implement total control and target management, collect background data and accumulate long-term monitoring data, so as to provide subsequent forecast and forecast environmental quality data.

In this paper, the "top-down" technical procedure [2-11] is used to determine sulfur dioxide in ambient air by ultraviolet fluorescence method as an example used to confirm the metrology traceability of the reliability of the online monitoring system for in-plant emissions SO_2, and a method to evaluate its uncertainty is given.

3. Statistical Methods and Cases of Online System SO_2

3.1. Top-Down Technology Overview

This technique focuses on evaluating the period precision of the data set under the control of the online system process [12-15]. Since the long-term errors of the system are directly related to the contribution sources of various factors, the systematic effects of these small drifts can be taken into account in the time variation sources. In addition, it is difficult to estimate the interaction effect between variables, but the implementation of top-down technology can be used to combine this effect to the maximum extent, which is proved by the independent identical distribution (i.i.d) of A* hypothesis test in Literature [1-11].

In this paper, Y method is regarded as the accepted benchmark, a generally reliable to verify the quantitative traceability of the reliability of on-line automatic systems (X method), and to evaluate its uncertainty.

Using a series of standard gas samples, based on the procedures of two measuring systems, calibrator (referred to as Y method) and monitor (referred to as X method), a series of horizontal measured values and period precision standard deviation $s_{R'}$, top-down site precision, were respectively given within the range of daily methods, and linear fitting relationship between $s_{R'}$ and level was established.

According to Reference 1, the approximate sum of squares (CSS) of residual-weighted methods were calculated successively: CSS_0, CSS_1 and CSS_2.

T test was used to select the desired CSS, and chi-square test showed that the selected CSS did not have sample matrix effect.

The i.i.d of *A** hypothesis test was conducted on the series residuals between them to determine whether the automatic monitoring system was under bias control. If the original hypothesis is true, the combined value of RXY is established and the uncertainty estimate of the system is given.

3.2. Calculation of Statistics

3.2.1. CSS_0

See Equation (1):

$$CSS_0 = \sum_i w_i \left(X_i - Y_i \right)^2 \tag{1}$$

Where,

$$w_i —— \frac{1}{s_{R'(Y)}^2 + s_{R'(X)}^2}$$

3.2.2. CSS_1

See Equation (2):

$$CSS_1 = \sum_i w_i \left[Y_i - \left(X_i + a \right) \right]^2 \tag{2}$$

Where,

$$w_i —— \frac{1}{s_{R'(Y)}^2 + s_{R'(X)}^2};$$

$$a —— \frac{\sum_i w_i Y_i}{\sum_i w_i} - \frac{\sum_i w_i X_i}{\sum_i w_i}.$$

3.2.3. CSS_2

See Equation (3-9):

$$\overline{X} = \frac{\sum_i w_i X_i}{\sum_i w_i} \tag{3}$$

$$\overline{Y} = \frac{\sum_i w_i Y_i}{\sum_i w_i} \tag{4}$$

Where,

$$w_i —— \frac{1}{s_{R'(Y)}^2 + b^2 s_{R'(X)}^2}.$$

$$x_i = X_i - \overline{X} \tag{5}$$

$$y_i = Y_i - \overline{Y} \tag{6}$$

$$b_0 = \frac{\sum w_i x_i y_i}{\sum w_i x_i^2 - \sum w_i^2 s_{R'(X)}^2 (y_i - b x_i)^2} \tag{7}$$

Note: Set b=1. If $|b-b_0|>0.001b$, replace b with b_0, recalculate w_i, $\overline{X}$ and $\overline{Y}$, x_i and y_i, as well as b0; Otherwise, the iteration stops.

$$CSS_2 = \sum_i w_i \left(y_i - b x_i \right)^2 \tag{8}$$

$$a = \overline{Y} - b\overline{X} \tag{9}$$

3.2.4. F and t Check for CSS Selection

The selection test of CSS is shown in Eq. (10) ~ (12):

$$F = \frac{(CSS_0 - CSS_2)/2}{CSS_2/(N-2)} \tag{10}$$

$$t_1 = \sqrt{\frac{CSS_0 - CSS_1}{CSS_2/(N-2)}} \tag{11}$$

$$t_2 = \sqrt{\frac{CSS_1 - CSS_2}{CSS_2/(N-2)}} \tag{12}$$

Note: If the calculated value of t_2 is larger than the quantile, CSS_2 is selected; otherwise, t_1 is calculated. If the calculated value of t_1 is larger than the quantile, select CSS_1; otherwise, select CSS_2.

3.2.5. χ2 and A* Inspection

The selected CSS values are compared with 95% quantiles of the χ2 distribution, where CSS_0 has N degrees of freedom; CSS_1 has N-1 degrees of freedom; CSS_2 has N-2 degrees of freedom. If the CSS value is less than the quantile, the hypothesis is valid.

Based on the selected CSS, the X_i prediction is used to calculate $\hat{Y}_i$, the following values εi of the data set according to Equation (13):

$$\varepsilon_i = \sqrt{w_i}\,(Y_i - \hat{Y}_i) \tag{13}$$

The ascending order ε_i is converted into v_i, see Equation (14):

$$v_i = (\varepsilon_i - \overline{\varepsilon}) / s_{\varepsilon} \tag{14}$$

It is converted vi into normal probability p_i value and substituted into equation (15) and equation (16):

$$A = -\frac{\sum_{i=1}^{n}(2i-1)[\ln(p_i)+\ln(1-p_{n+1-i})]}{n} - n \quad (15)$$

$$A^* = A(1+\frac{0.75}{n}+\frac{2.25}{n^2}) \quad (16)$$

If the calculated value is < 0.752, it indicates that the series values are normally distributed; otherwise, the null hypothesis is rejected.

3.2.6. Uncertainty Evaluation

The uncertainty evaluation of automatic monitoring system is shown in Equation (17):

$$U = 0.722\sqrt{\frac{(R_Y^{'})^2 + b^2(R_X^{'})^2}{2}} \quad (17)$$

Where,

b -- If CSS_0 and CSS_1 are selected, b=1.

$\hat{Y}$ -- The bias correction result of X method. When measuring the same sample, the interval covers $\hat{Y} \pm R_{X\hat{Y}}$ the result of Y method with 95% probability.

4. Application Cases

4.1. Precision Measurement During the Period

The volume percentage of daily contact range of gaseous pollutant sulfur dioxide is 100 ×10-9 ~ 400×10-9. The automatic monitoring system of an environmental monitoring station is based on Y method and X method, ultraviolet (UV) fluorescence method, the same series of standard gases, based on the statistical procedures. The quality level of the diluted concentration was tested by continuous tracking.

The standard gas concentration selected in this case is 52.1×10-6, and it is diluted by Y method (model: SABIO 4010). The volume percentage concentration after dilution is 100×10-9, 200 ×10-9, 300 ×10-9 and 400×10-9, and then combined with X method (model: EC 9850), contemporaneous precision measurements of diluted horizontal concentrations were made.

Table 1 shows the repetition of measurement times N=27 of the two methods respectively and under $s_{R'}$, top-down site precision, level. In view of the significant difference between $s_{R'}$ under level, linear fitting relation between $s_{R'}$ and level is given.

4.2. Fitting of CSS_0 and CSS_1

According to equations (1) and (2), statistics in Table 1 can be calculated as follows:

$$\sum_i w_i = \frac{1}{s_{R'(Y)}^2 + s_{R'(X)}^2} = 27.93,$$

$$CSS_0 = \sum_i w_i (X_i - Y_i)^2 = 30.98,$$

$$a = \frac{\sum_i w_i Y_i}{\sum_i w_i} - \frac{\sum_i w_i X_i}{\sum_i w_i} = 0.0176,$$

$$CSS_1 = \sum_i w_i [Y_i - (X_i + a)]^2 = 30.97.$$

Table 1. *CSS_0 and CSS_1 fitting under different level measurement estimates s(R').*

lot	*Y* method		*X* method		Calculation of CSS_0 and CSS_1				
	Y_i	$s_{R'(Y)}$	X_i	$s_{R'(X)}$	w_i	$w_i(Y_i-X_i)^2$	w_iY_i	w_iX_i	$w_i(Y_i-X_i-a)^2$
1	101	0.270	101	0.593	2.355	0.212	237.68	236.97	0.188
2	101	0.269	100	0.592	2.364	0.378	238.09	237.15	0.346
3	100	0.268	99.8	0.590	2.380	0.595	238.69	237.50	0.554
4	100	0.268	99.5	0.589	2.388	1.170	239.26	237.59	1.112
5	101	0.270	101	0.594	2.349	0.000	236.97	236.97	0.001
6	99.9	0.267	101	0.593	2.362	1.512	235.97	237.86	1.579
7	99.6	0.266	99.2	0.588	2.400	0.384	239.08	238.12	0.351
8	99.1	0.264	99.6	0.590	2.395	0.599	237.38	238.58	0.642
9	201	0.567	201	0.950	0.817	0.400	163.85	164.42	0.421
10	201	0.569	201	0.948	0.818	0.401	164.74	164.17	0.381
11	200	0.564	201	0.951	0.819	2.956	163.36	164.91	3.011
12	200	0.566	199	0.941	0.829	1.625	165.98	164.82	1.584
13	199	0.563	199	0.943	0.829	0.008	165.28	165.20	0.006
14	199	0.561	201	0.949	0.823	4.354	163.48	165.37	4.421
15	302	0.868	302	1.307	0.406	0.000	122.60	122.60	0.000
16	301	0.865	301	1.304	0.409	0.037	122.84	122.97	0.041
17	299	0.861	299	1.298	0.412	0.004	123.41	123.37	0.003
18	299	0.859	298	1.294	0.415	0.066	123.80	123.64	0.061
19	301	0.867	302	1.306	0.407	0.037	122.62	122.74	0.041
20	402	1.166	402	1.663	0.242	0.087	97.38	97.53	0.092
21	398	1.154	401	1.660	0.245	3.351	97.32	98.22	3.383

lot	*Y* method		*X* method		Calculation of CSS_0 and CSS_1				
	Y_i	$s_{R'(Y)}$	X_i	$s_{R'(X)}$	w_i	$w_i(Y_i-X_i)^2$	w_iY_i	w_iX_i	$w_i(Y_i-X_i-a)^2$
22	398	1.156	401	1.658	0.245	1.530	97.51	98.12	1.552
23	402	1.168	398	1.647	0.245	5.419	98.71	97.56	5.378
24	402	1.166	400	1.655	0.244	0.791	98.03	97.59	0.775
25	402	1.168	398	1.648	0.245	4.120	98.58	97.57	4.085
26	399	1.157	401	1.657	0.245	0.884	97.62	98.09	0.901
27	402	1.167	402	1.664	0.242	0.061	97.35	97.47	0.065
	$s_{R'(Y)}$=0.267+0.003(Y_i-100)		$s_{R'(X)}$=0.591+0.004(X_i-100)		27.93	30.98	4287.6	4287.1	30.97
							153.50	153.48	0.018

4.3. Fitting of CSS2

According to equations (3) ~ (9), statistics in Table 2 can be calculated as follows:

$$\sum w_i y_i x_i = 228146.7\ ,\ \sum w_i x_i^2 = 228312.2\ ,\ \sum w_i^2 s_{R'(X)}^2 (y_i - bx_i)^2 = 22.18\ ,$$

$$b_0 = \frac{\sum w_i x_i y_i}{\sum w_i x_i^2 - \sum w_i^2 s_{R'(X)}^2 (y_i - bx_i)^2} = 0.9994$$

Because | b - b_0 | = | | 1-0.999-4 = 0.000 6 < 0.001 b, don't need to continue the iteration.

$$CSS_2 = \sum_i w_i \left(y_i - bx_i\right)^2 = 22.18\ ,\ a = \overline{Y} - b\overline{X} = 153.50 - \left(0.9994 \times 153.48\right) = 0.1140\ .$$

Table 2. *CSS_2 fitting under different level measurement estimates s(R').*

lot	Y_i	X_i	w_i	w_iY_i	w_iX_i	y_i	x_i	$w_iy_ix_i$	$w_ix_i^2$	$[w_is_{R'(X)}(y_i-bx_i)]^2$
1	101	101	2.355	237.68	236.97	-52.60	-52.88	6 552.8	6 587.9	0.155
2	101	100	2.364	238.09	237.15	-52.80	-53.18	6 639.7	6 687.8	0.286
3	100	99.8	2.380	238.69	237.50	-53.20	-53.68	6 796.9	6 858.5	0.459
4	100	99.5	2.389	239.26	237.59	-53.30	-53.98	6 870.9	6 958.8	0.922
5	101	101	2.349	236.97	236.97	-52.60	-52.58	6 496.3	6 494.1	0.001
6	99.9	101	2.362	235.97	237.86	-53.60	-52.78	6 683.3	6 581.3	1.314
7	99.6	99.2	2.400	239.08	238.12	-53.90	-54.28	7 023.8	7 073.7	0.292
8	99.1	99.6	2.395	237.38	238.58	-54.40	-53.88	7 021.9	6 955.1	0.534
9	201	201	0.817	163.85	164.42	47.10	47.82	1 839.5	1 867.5	0.310
10	201	201	0.818	164.74	164.17	47.80	47.12	1 843.0	1 816.7	0.280
11	200	201	0.819	163.36	164.91	46.00	47.92	1 804.7	1 880.0	2.228
12	200	199	0.829	165.98	164.82	46.70	45.32	1 754.4	1 702.5	1.164
13	199	199	0.829	165.28	165.20	45.80	45.72	1 736.3	1 733.2	0.004
14	199	201	0.823	163.48	165.37	45.10	47.42	1 760.1	1 850.6	3.277
15	302	302	0.406	122.60	122.60	148.30	148.32	8 935.3	8 936.4	0.000
16	301	301	0.409	122.84	122.97	147.10	147.42	8 861.5	8 880.7	0.029
17	299	299	0.412	123.41	123.37	145.90	145.82	8 768.9	8 763.9	0.002
18	299	298	0.415	123.80	123.64	145.10	144.72	8 706.0	8 683.0	0.042
19	301	302	0.407	122.62	122.74	147.80	148.12	8 909.2	8 928.4	0.028
20	402	402	0.242	97.38	97.53	248.10	248.72	14 962.9	15 000.1	0.062
21	398	401	0.245	97.32	98.22	244.10	247.82	14 805.7	15 031.2	2.280
22	398	401	0.245	97.51	98.12	244.80	247.32	14 822.1	14 974.5	1.044
23	402	398	0.245	98.71	97.56	248.90	244.22	14 910.4	14 629.9	3.578
24	402	400	0.244	98.03	97.59	248.20	246.42	14 925.9	14 818.7	0.518
25	402	398	0.245	98.58	97.57	248.70	244.62	14 910.8	14 666.0	2.720
26	399	401	0.245	97.62	98.09	245.10	247.02	14 827.5	14 943.5	0.605
27	402	402	0.242	97.35	97.47	248.40	248.92	14 976.9	15 008.1	0.043
			27.93	4 287.6	4 287.1			228 146.7	228 312.2	22.18
				153.50	153.48	*a*=0.114 0				*b*=0.999 4

4.4. Statistical Verification of CSS Selection

Given as $CSS_0 = 30.98, CSS_1 = 30.97, CSS_2 = 22.18$,

Then, according to Equation (10), the calculation is as follows:

$F=\frac{(CSS_0-CSS_2)/2}{CSS_2/(N-2)}=9.919>F_{0.95}(2.25)=3.385$, it is shown that CSS bias correction can improve the expected consistency between methods.

According to Equation (12), $t_2=\sqrt{\frac{CSS_1-CSS_2}{CSS_2/(N-2)}}=9.92>t_{0.975}(25)=2.06$, Therefore, CSS_2 is selected, i.e $\hat{Y}=0.1140+0.9994X$, as $\chi^2=CSS_2=22.18<\chi^2_{0.99}(25)=37.65$, i.i.d of *A** test and uncertainty evaluation can be carried out.

4.5. A Testing and Uncertainty Evaluation*

Based on $\hat{Y}=0.1140+0.9994X$, through X_i to predict the fit value $\hat{Y}_i$. According to equations (13) ~ (16), normality test was conducted for series residuals εi between methods.

For series ε_i, *A**=0.246 <*A**(0.05)=0.752, indicating that the measured results of the two methods are normally distributed. See Table 3 for details.

Table 3. *Statistical summary of A*tests.*

lot	Y_i	$s_{R'(Y)}$	X_i	$s_{R'(X)}$	w_i	$\sqrt{w_i}$	$\hat{Y}$	ε_i	p_i asce.	p_i desce.	The ith term of A
1	101	0.270	101	0.593	2.358	1.535	100.65	0.378	0.028	0.986	-7.83
2	101	0.269	100	0.592	2.367	1.538	100.35	0.533	0.053	0.973	-19.64
3	100	0.268	99.8	0.590	2.382	1.543	99.85	0.688	0.059	0.881	-24.83
4	100	0.268	99.5	0.589	2.390	1.546	99.55	0.998	0.115	0.822	-27.19
5	101	0.270	101	0.594	2.351	1.533	100.95	-0.082	0.142	0.810	-32.54
6	99.9	0.267	101	0.593	2.364	1.538	100.75	-1.313	0.212	0.738	-31.80
7	99.6	0.266	99.2	0.588	2.403	1.550	99.25	0.536	0.217	0.723	-36.55
8	99.1	0.264	99.6	0.590	2.398	1.548	99.65	-0.858	0.285	0.691	-36.44
9	201	0.567	201	0.950	0.817	0.904	201.29	-0.627	0.418	0.690	-34.73
10	201	0.569	201	0.948	0.819	0.905	200.59	0.639	0.436	0.638	-35.10
11	200	0.564	201	0.951	0.820	0.905	201.39	-1.714	0.448	0.611	-36.67
12	200	0.566	199	0.941	0.830	0.911	198.79	1.280	0.448	0.541	-36.39
13	199	0.563	199	0.943	0.830	0.911	199.19	0.096	0.472	0.538	-38.03
14	199	0.561	201	0.949	0.824	0.908	200.89	-2.082	0.518	0.518	-37.47
15	302	0.868	302	1.307	0.407	0.638	301.73	0.043	0.538	0.472	-36.55
16	301	0.865	301	1.304	0.409	0.639	300.83	-0.149	0.541	0.448	-37.46
17	299	0.861	299	1.298	0.412	0.642	299.23	0.106	0.611	0.448	-35.88
18	299	0.859	298	1.294	0.415	0.644	298.14	0.299	0.638	0.436	-35.75
19	301	0.867	302	1.306	0.407	0.638	301.53	-0.149	0.690	0.418	-33.76
20	402	1.166	402	1.663	0.243	0.493	402.07	-0.233	0.691	0.285	-27.50
21	398	1.154	401	1.660	0.245	0.495	401.17	-1.768	0.723	0.217	-23.33
22	398	1.156	401	1.658	0.245	0.495	400.67	-1.175	0.738	0.212	-23.30
23	402	1.168	398	1.647	0.245	0.495	397.58	2.390	0.810	0.142	-16.34
24	402	1.166	400	1.655	0.244	0.494	399.77	0.952	0.822	0.115	-14.99
25	402	1.168	398	1.648	0.245	0.495	397.98	2.092	0.881	0.059	-9.16
26	399	1.157	401	1.657	0.245	0.495	400.37	-0.878	0.973	0.053	-4.17
27	402	1.167	402	1.664	0.242	0.492	402.27	-0.183	0.986	0.028	-2.27
											A*=0.246

As table 1 give $s_{R'(Y)}$= 0.267+0.003(*Y*-100) and $s_{R'(X)}$= 0.591+0.004(*X*-100), then R_Y'=2.77[0.267+0.003(*Y*-100)]=0.008(*Y*-11.7), R_X'=2.77[0.591+0.004(*X*-100)]=0.011(*X*-48.1). Using Formula (17), the estimated value of the extended uncertainty (U) is given:

$$U=0.722\sqrt{\frac{(R_Y')^2+b^2(R_X')^2}{2}}=0.722\sqrt{\frac{[0.008(Y-11.7)]^2+0.9994^2\times[0.011(X-48.1)]^2}{2}}=\sqrt{0.00002(Y-11.7)^2+0.00003(X-48.1)^2}$$

This fitted variation values is smaller than the horizontal technical indicators, in the technical performance indicators of environmental air automatic monitoring Instruments.

5. Results and Discussion

5.1. Period Precision and Statistical Control

The long-term error of automatic monitoring system is directly related to random factors such as time interval, personnel operation, gas dilution and instrument calibration, which cannot be detected under repeated test conditions, but the period precision test covers the source of variation as far as possible and avoids complex calculation of correlation factors.

From the perspective of mathematical statistics, the main cause of statistical runaway is systematic error, where the mass characteristics do not obey the normal distribution. Although data accuracy depends on the control of systematic errors, it also depends to some extent on random errors. In fact, in any test, two kinds of errors exist simultaneously. In daily life, unmastered systematic errors, or those that can be mastered but are too complex and regular, are often treated as random errors after passing *A** normality test, and treated technically appropriately through the bias correction in this paper.

It is not easy to ensure statistical control. To test whether the quality characteristics of the process obey the normal distribution without the influence of time and space, in addition to relying on professional technical knowledge and previous experience, the *A** random effect hypothesis test and data transformation of heteroscedasticity effect in this paper can also be used to study and identify the stability of the quality and its change trend.

In the actual test, α and β error rates exist simultaneously, so it is a complicated thing to evaluate the reliability of test results. According to the above analysis, a test with a small amount of bias and a small standard deviation is likely to be more reliable than a test with 0 bias and a large standard deviation. Because there is not always an opportunity to perform a large number of measurements, one cannot always rely on the reliability of the mean. One should strive to be reliable every single test. Therefore, it is necessary to strengthen the predictive control of random factors, narrow the fluctuation range of test quality, and finally achieve the purpose of improving the stability and precision of automatic monitoring system.

5.2. CSS Bias Correction and Hypothesis Testing

Due to the heterogeneity of observation data, heteroscedasticity will occur when the regression of different sample means in a covariable is processed. In this case, only weighted regression of CSS fitting is helpful for analysis.

The F test of CSS is designed to focus on the degree of agreement between actual and theoretical distributions. In view of the general set of alternative hypotheses, F ratio is mostly set by unilateral test, that is, dividing by CSS_2 for normalization, to check whether CSS function can satisfy the linear relationship between Y and X. If the values of these numerators exceed those of their denominators, it is entirely due to the real difference between the population means, that is, the source of variation of the difference between the standard gases is considered to be greater than the common CSS_2 of the population.

The t-test of CSS means that the null hypothesis is that the population average of the difference is zero, and the alternative hypothesis is that the average cannot be zero. If t_2 rejects the null hypothesis and believes that the difference is too large to be equal to zero, CSS_2 should be selected. If t_2 accepts the null hypothesis, CSS_1 will be compared; if t_1 rejects the null hypothesis, CSS_1 will be compared; if t_1 rejects the null hypothesis, CSS_2 will be compared.

The $\chi 2$ test of CSS can be directly compared with the $\chi 2$ degree of freedom. If there is no significant difference, there is no sample matrix effect. If matrix effects exist, they can be remedied by *A** testing, and if the original hypothesis is valid, the treated random effects are incorporated into the uncertainty.

5.3. Follow-up Monitoring and Uncertainty Update

The precision and bias performance parameters of the automatic monitoring system need to be updated at any time to be adjusted and improved. Prolonged repetition of testing in the context of periodic precision measurement can result in changes in equipment and environment, which are entirely dependent on the degree of indoor control. With the passage of time, there is also random fluctuation between batches, that is, reagent degradation, instrument drift and adjustment, temperature change, personnel operation and other effects are summed up as repeatable variable factors, which constitute random fluctuation effect between batches of samples. If the indoor wants to know such variation, it is necessary to carry out targeted updating and tracking research of system performance control parameters, so as to provide better variation information and necessary tolerance limits, which is conducive to the improvement of system performance trend.

This study pursues economy and emphasizes the unified principle of risk, benefit and cost. As long as the index requirements of product quality can be satisfied, the error risk generated under the use conditions can be controlled within a certain limit after the automatic monitoring system eliminates the systematic error. The uncertainty must be evaluated repeatedly over time. When the uncertainty of long-term monitoring reaches constant, it is helpful to automatically monitor the bias correction of the system. Through multi-batch testing, this study increased the degree of freedom, accurately evaluated the measured uncertainty, and saved the control costs and resources invested in previous repeated experiments.

6. Conclusion

(1) In this study, under the statistical controlled analysis of the precision measurement during the period, the uncertainty assessment brought by the bias correction fitting line was considered, so as to combine the influence of various variation factors to the maximum extent, and keep the error sources independent from each other, so as to avoid the complicated calculation of correlation, which is helpful for the quality assurance and quantification traceability of the automatic monitoring online analysis and measurement system.

(2) This study puts forward a new conclusion: the uncertainty contribution of the automatic monitoring online analysis and measurement system should be composed of two parts: standard gas and the variation of bias correction linear fitting. This conclusion can minimize the risk caused by the wrong assessment of uncertainty.

(3) This study regards "measurement" as "process", especially emphasizing that not only the final statistical analysis of data, but also the principles proposed in this study should be fully applied at the

beginning of data collection. Only when the measurement is under statistical control, the uncertainty assessed is effective, and the hypothesis of scientific and accurate expected conclusions can be reached.

(4) It is suggested in this study that the established bias $\hat{Y} = 0.1140 + 0.9994X$ should be continuously tracked and monitored at any time, and the evaluated bias should be adjusted and $U = \sqrt{0.00002(Y - 11.7)^2 + 0.00003(X - 48.1)^2}$ improved constantly, which will help further optimize the quality objectives of the automatic monitoring online system.

(5) This study can dynamically track and monitor the periodic changes of the operation of the automatic monitoring online system, verify whether the bias correction relationship and uncertainty assessed represent the actual performance of the daily measurement system, so as to detect the potential influencing factors or the hidden dangers prone to occur at any time, and propose corresponding improvement measures.

References

[1] ASTM D6708-2021: Standard Practice for Statistical Assessment and Improvement of Expected Agreement Between Two Test Methods that Purport to Measure the Same Property of Material, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[2] ASTM D6299-2022el: Standard Practice for Applying Statistical Quality Assurance Techniques to Evaluate Analytical Measurement System Performances, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[3] ASTM D6617-2021: Standard Practice for Laboratory Bias Detection Using Single Test Result from Standard Material, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[4] ASTM D6792-2023a: Standard Practice for Quality Management Systems in Petroleum Products, Liquid Fuels, and Lubricants Testing Laboratories, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[5] CNAS-GL016-2020: Guidance and illustration on Uncertainty Estimation in Physical and Chemical Testing in the Field of Petroleum and Petrochemicals, China National Accreditation Service for Conformity Assessment, Beijing, www.cnas.org.cn

[6] Yang Shuo, Pan ZhiQiang Pan, Wang Dou Wen. *Uncertainty Evaluation of Automatic Monitoring System for Fine Particulate Matter in Ambient Air*, International Journal of Mechanical Engineering and Applications, Vol. 9, No. 6, 2021, pp. 95-98.

[7] Zeng Xing-yu, et al. *Evaluation of Uncertainty in Determination of Total Organic Carbon in Concentrated Seawater by Using Control Chart in Top-down*. Contemporary Chemical Industry, 2021, 50 (1), pp. 95-1252.

[8] Wang Qiang, et al. *Uncertainty evaluation of determination of components in natural gas by using Top-down method*. Chemical Engineering of Oil $ Gas, 2020, 48 (3), pp. 98-103.

[9] Gao Huan, et al. *Measurement Uncertainty for Unwashed Gum Content of Gasoline by Top-down Method*. Guang Dong Chemical, 2021, 48 (9), pp. 263-267.

[10] Cao Yiqun, et al. *Estimation of Measurement Uncertainty of AFB by Top-down*. Farm Products Processing, 2020, 1 (1), pp. 51-52.

[11] Zhang Jifei, et al. *Estimation of Measurement Uncertainty for Determination of Lead in Plastic Toys by X-ray fluorescence spectrometer based on Top-down method*. Quality Safety Inspection and Testing, 2021, 31 (2), pp. 6-8.

[12] ASTM D3764-2022: Standard Practice for Validation of the Performance of Process Stream Analyzer Systems, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[13] ASTM D6122-2022: Standard Practice for Validation of the Performance of Multivariate Online, At-Line, Field and Laboratory Infrared Spectrophotometer, and Raman Spectrometer Based Analyzer Systems, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[14] ASTM D7235-2021a: Standard Guide for Establishing a Linear Correlation Relationship Between Analyzer and Primary Test Method Results Using Relevant ASTM Standard Practices, ASTM International, West Conshohocken, PA, 2010, www.astm.org.

[15] ASTM D7808-2022: Standard Practice for Determining the Site Precision of a Process Stream Analyzer on Process Stream Material, ASTM International, West Conshohocken, PA, 2010, www.astm.org.