

Article

A Flexible Data Evaluation System for Improving the Quality and Efficiency of Laboratory Analysis and Testing

Yonghui Tu ¹, Haoye Tang ¹, Hua Gong ¹ and Wenyu Hu ^{2,*}

¹ State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

² Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

* Correspondence: wyhu@issas.ac.cn

Abstract: In a chemical analysis laboratory, sample detection via most analytical devices obtains raw data and processes it to validate data reports, including raw data filtering, editing, effectiveness evaluation, error correction, etc. This process is usually carried out manually by analysts. When the sample detection volume is large, the data processing involved becomes time-consuming and laborious, and manual errors may be introduced. In addition, analytical laboratories typically use a variety of analytical devices with different measurement principles, leading to the use of various heterogeneous control software systems from different vendors with different export data formats. Different formats introduce difficulties to laboratory automation. This paper proposes a modular data evaluation system that uses a global unified management and maintenance mode that can automatically filter data, evaluate quality, generate valid reports, and distribute reports. This modular software design concept allows the proposed system to be applied to different analytical devices; its integration into existing laboratory information management systems (LIMS) could maximise automation and improve the analysis and testing quality and efficiency in a chemical analysis laboratory, while meeting the analysis and testing requirements.

Keywords: chemical analysis; sample detection; raw data; heterogeneous system; data formats; evaluate quality; modular software; LIMS



Citation: Tu, Y.; Tang, H.; Gong, H.; Hu, W. A Flexible Data Evaluation System for Improving the Quality and Efficiency of Laboratory Analysis and Testing. *Information* **2022**, *13*, 424. <https://doi.org/10.3390/info13090424>

Academic Editor:
Aneta Poniszewska-Maranda

Received: 4 August 2022
Accepted: 29 August 2022
Published: 8 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Data evaluation includes comprehensive data processing, investigation, as well as evaluation, storage, and output from respective data applications to improve data reliability and effectiveness. Data evaluation also provides a more favourable basis for decision-making [1]. Traditional chemical analysis laboratories usually undertake many tasks, and many sample detections produce large amounts of data. Taking the laboratory where the authors work as an example, the continuous flow analysis (CFA) analyses about 15,000 samples, and the HPLC-ICP-MS, ICP-AES analyses about 45,000 samples, every year. Experimental analyses of these samples would generate a large amount of data every year. Therefore, the quality and efficiency of data evaluation is extremely important. The traditional processing method involves manually processing all relevant tasks. After detection, analysts export all the original experimental data from the analytical device software system, filter out the corresponding item data according to customer needs, and sort the data to evaluate its effectiveness. Sometimes, statistical analyses and validation calculations are needed, such as the determination of precision and detection and quantification limits. Finally, effective item data are required to formulate valid data reports for distribution. In addition, the analysts also need to evaluate the experimental process according to the relevant information obtained from the original data, in order to judge the stability and reliability of the analytical device state. These tasks are usually completed manually by laboratory personnel. When the number of sample detections is large and processing time

is critical, this traditional data evaluation method will become particularly time-consuming, laborious, inefficient, and may also introduce manual errors [2,3]. Therefore, effective data evaluation tools are becoming increasingly important in daily sample testing at chemical analysis laboratories.

Taking the interval flow analyser as an example, which is an inorganic analytical device that is supervised by one of the authors, the original data file generated by the device after an experiment contains the concentration value of the analysis item and other information, such as signal strength, peak time, and peak identification. However, this information is not the output required by customers in the final data report. Therefore, analysts must manually screen and separate the original data files in order to obtain the concentration values of the analysis items. On the other hand, for experimental analysts, the rich data are usually necessary to judge the stability and reliability of the state of the analytical device according to the data related to the detection signal, which requires relevant information such as the original signal strength of the sample, peak time, baseline signal strength, quality control signal strength, sensitivity correction, etc., from the original data file [4,5].

According to the authors' years of work experience, in addition to long-term communication with other peers and customers, for the same original data produced by analytical devices, it is often necessary to output different results according to different target needs. Thus, for the experimental data output, analysts often need to carry out processing. For example, analysts use statistical methods and combine quality control results to evaluate the effectiveness of the experimental data and correct the errors, marking the data that do not meet the quality requirements, and retesting the samples [6]. Therefore, in any case, it is difficult for us to directly obtain the final data products from the device system after detection. For most analytical devices, there is also a last-mile problem between the device system itself and customers' particular needs for the experimental data. Using a unified system to meet these differentiated needs is a complex challenge [7].

Generally speaking, different analytical device manufacturers provide different software systems, and the software also varies depending on the type of analytical device that is involved. Even for the same type of device, its software functions may vary from manufacturer to manufacturer, and the formats of exported data also depend on the manufacturer (such as table layout); this means that the data and table styles provided by different software solutions are inconsistent with each other [8]. These factors further increase the difficulty and complexity of manual completion.

In order to meet these challenges and improve work quality and efficiency, various studies have been completed. Zhu, B.Y. et al. summarised the applications of artificial intelligence in contemporary chemical research, and pointed out that artificial intelligence can mine correlations in massive experimental data generated in chemical experiments, help chemists make reasonable analyses and predictions, and consequently accelerate the process of chemical research [9]. Emerson, J. et al. pointed out that multivariate data analysis (MVDA) has been successfully applied to various applications in the chemical and biochemical industries, including bioprocess monitoring, identification of critical process parameters, assessment of process variability, comparability during process development, scale-up, and technology transfer [10]. Qi, Y.L. et al. introduced data processing in Fourier transform ion cyclotron resonance mass spectrometry [11]. Urban, J. et al. described the currently available methods and techniques commonly used in HPLC-MS for processing and analysis, as well as the main types of processing and most popular methods [12]. These studies often focused on the purpose of experimental research or the method of analysing experimental data, but did not address the process of obtaining the experimental data. In addition, the high-throughput screening method can generate a large amount of data in a short time; furthermore, this method is used not only in the field of biotechnology and bio-screening [13,14], but also in the fields of analytical chemistry (such as catalyst research [15]), clinical studies, and toxicology [8,16]. Moreover, laboratory information management systems have also been applied to the organisation and storage of proteomic

data [17]. In the data evaluation of these fields, many individual software solutions have been developed for flexible applications.

The authors of this paper have been engaged in analytical devices management and sample analysis in a traditional chemical analysis laboratory affiliated with the Institute of Soil Science, the Chinese Academy of Sciences, for more than ten years. This is a traditional modern chemical analysis laboratory that uses various analytical devices. The analysis and detection business covers many fields, such as soil, water, and plants, and has a perfect quality assurance system and a standardised service management process. Customers are mainly internal or external scientific research personnel at the Institute, and the sources of samples are mainly scientific research projects of customers, including soil samples and solution samples. The devices supervised by the authors include the CFA, HPLC-ICP-MS, ICP-AES, and other inorganic analysis devices. This paper is the authors' experience and technical summary regarding data processing and evaluation in sample analysis work conducted over many years. The purpose of this paper is to develop a special data evaluation system for the interval flow analyser of inorganic analytical devices, aiming at a series of business problems faced by laboratory devices after analysis and testing, such as original data screening, editing, effectiveness evaluation, error correction, etc.; the objective is to meet analysis and testing needs, and improve quality and efficiency. At the same time, the idea of combining the business process with the development and application of the proposed data evaluation system is quite representative. The proposed system is also applicable to the sample analysis business of many other devices in chemical analysis laboratories.

The main contributions of this paper are as follows:

1. This paper combs and analyses the data evaluation work in the analysis and detection business of traditional chemical analysis laboratories and points out the common problems and risks in the traditional data evaluation business.
2. Taking the inorganic analytical device for the interval flow analyser as an example, a data evaluation system is developed, which realizes automatic data screening, quality evaluation, data management and distribution, integrates with the existing LIMS, provides the maximum automation effect, and improves the quality and efficiency of the analysis and testing business.
3. The idea of modular design makes it easy for a data evaluation system to be partially or wholly extended to different analysis systems produced by different analysis device manufacturers. This provides a reference method for data evaluation work in a chemical analysis laboratory, as well as a reference for improving the quality and efficiency of analysis and testing.

The rest of this paper is organised as follows: Section 2 introduces the overall architecture design of the system, including the functions and composition of each module, as well as the relationships between them. Section 3 describes the functions of each system software module, the problems that are solved, and their specific development design. Section 4 verifies the application and effect of each system module according to the interval flow analyser analysis and detection processes. The last section summarises the application effect of the data evaluation system developed in this paper specifically with regards to analytical devices, and discusses the significance and implementation method of extending the design idea of this paper to different analytical systems produced by different analytical device manufacturers.

2. Overall Structure of the System

Generally, analysis and detection processes in chemical analysis laboratories include chemical analysis and instrumental analysis. This paper studies instrumental analysis, which generally involves relatively complex or special instruments and equipment to obtain chemical compositions, components content, chemical structures, and other information of substances. Instruments measure the parameters and changes in various physical or physicochemical properties, and use relevant software to conduct qualitative and quantita-

tive analyses of the results. The measurement results are stored in an original data file [18]. In most cases, these data can be exported in the form of Microsoft Excel or CSV files, and external software modules can be used for automatic data evaluation.

This paper uses interval flow analyser data as an example. According to its business process that uses an object-oriented programming language, a special data evaluation system was developed to assist business processes and improve business quality and efficiency. The overall structural diagram of the system is shown in Figure 1.

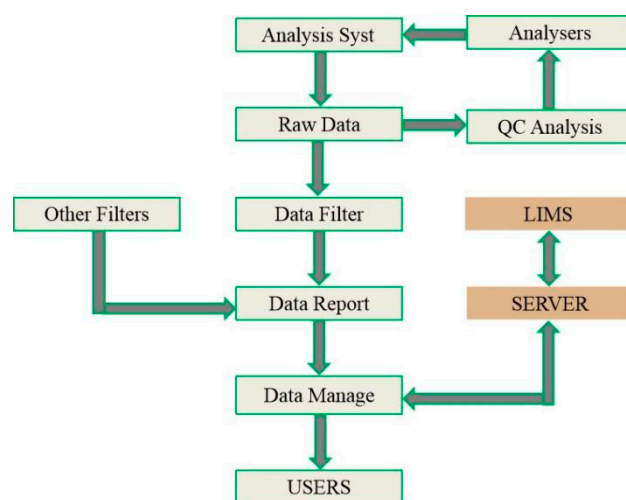


Figure 1. Overall structure of the system.

In the development of this system, we divide the system into four modules: data filtering; quality control analysis; valid data reports; and data management and distribution [19,20]. Data filtering and valid data reports are completed locally at the client, which are customer-oriented. Quality control analysis refers to the visual expression of the experimental quality by the analysts through the rich data of each analysis item, and the output results are oriented to the analysts. Finally, according to our business process, we need to distribute the standard reports to users and upload them to our LIMS for management. In the past, reporting the data documents was carried out manually, one by one; this is cumbersome, time-consuming, and inefficient. Generally, data document reporting processing for four businesses requires a total of 40 page clicks. The speed of the whole process depends on the manual proficiency of the experimenters, as well as the computer system and broadband resources. In addition, when the data reporting process is implemented locally, the local computer resources of the experimenters will be occupied in real time, and the processing must communicate with the corresponding business modules of the LIMS. The access of the LIMS system and the transmission of data documents will occupy the local bandwidth resources of the experimenters. Therefore, data management and distribution adopted the C/S architecture and were completed by the client, in collaboration with the local server to achieve efficient, low-latency reporting of data documents, along with persisting and high-reliability storage management and maintenance. The report documents are centrally archived locally, while the server is fully retained for backup. The server adopts Elastic Compute Service (ECS) provided by Alibaba Cloud, which has the characteristics of low latency, persistence, and high reliability [21].

3. Modular Software Development

3.1. Data Filter

The continuous flow analyser is based on the design by J. Ruzicka and E.H. Hansen, who proposed the concept of flow injection and designed an analytical instrument [22]. Based on continuous flow analysis (CFA), also called segmented flow analysis (SFA), pump pipes with different diameters are compressed through a peristaltic pump; the reaction

reagent and the sample to be tested are injected into a closed and continuous flow carrier in proportion. A colour reaction is produced in the chemical reaction unit, its signal value in the detector is measured, and the concentration of the sample to be tested is determined according to the standard curve method. The continuous flow analyser can be used to detect the content of soluble nitrogen and phosphorus in soil and water. It is an important instrument to support scientific research tasks related to the soil nitrogen cycle and global change, as well as nitrogen migration and transformation [23,24]. Taking the detection of total nitrogen (TN) as an example, the data file generated by the flow analyser after the experiment contains not only the concentration value of the analysis item, but also a great deal of other information, as shown in Table 1.

Table 1. Raw data generated by the CFA after detecting TN.

Position	Type	Identity	Ext.Dil	Weight	Pre.Dil	TN mg/L	Height TN	Corr.Ht TN	Flag TN	Time TN
WT	IW	Initial Wash	1	1	1	0	38,868	0	IW	805
ST1	T	Tracer	1	1	1	9.09	41,053	2204	N	1298
ST1	D	Drift	1	1	1	9.47	41,126	2296	N	1482
Wt	W	Wash	1	1	1	0	38,811	0	N	1646
Wt	S1	Standard 1	1	1	1	−0.01	38,810	−1	A	1846
E51	S2	Standard 2	1	1	1	3.17	39,605	769	N	2065
E52	S3	Standard 3	1	1	1	6.08	40,348	1474	A	2222
E53	S4	Standard 4	1	1	1	8.98	41,105	2178	A	2373
E54	S5	Standard 5	1	1	1	12.24	41,968	2967	N	2544
ST1	D	Drift	1	1	1	10.17	41,279	2465	N	2719
Wt	W	Wash	1	1	1	0	38,814	0	N	2904
D1	U	1	5	1	1	8.59	39,091	260	A	3077
D2	U	2	5	1	1	8.07	39,071	244	A	3258
D3	U	3	5	1	1	8.6	39,080	261	A	3808

Customers only need to obtain the concentration value information of the sample; most of the other field data are not necessary. In contrast, the experimenters must judge the stability and reliability of the experimental instrument according to the multidata information supplied by the detection signals. Therefore, even if the unnecessary data are eliminated, a single analysis item will output up to seven fields and seven types of data. If fields such as sample position, type, and number are added, more information will be output; if nitrate nitrogen, nitrite nitrogen, ammonia nitrogen, and total nitrogen are analysed at the same time, up to 35 fields and 7 types of data will appear. Usually after experimental completion, experimenters can only use the traditional manual method to process these data and generally, there is no problem in doing so. However, when we strive to improve experimental quality and efficiency, especially when the volume of sample analyses is large and time is limited, the traditional data evaluation method will become particularly time-consuming and laborious, and human errors may also be introduced. Therefore, we developed a data filtering module to realize the automatic filtering processing of the target data. According to the set effective value and limit of detection (LOD), the sample data information can be screened out, and the necessary data format processing can be carried out. The samples beyond the effective value can be marked, and the quality control data can be listed separately in order to make a preliminary judgement on the quality of the experimental analysis data. Figures 2 and 3 show the interface of the filtering module and the basic workflow of data filtering.

The module interface contains all of the analysis items of the CFA. During data filtering, we check the analysis items we want in order to extract according to need. The effective value and LOD give the maximum and minimum values of the effective value of the item, respectively. Values beyond this range will be regarded as abnormal and specially marked in the results. The business delegation from whence these data were processed will be listed under the order number. Generally, different analytical device manufacturers provide different software systems, and even the software functions may vary according

to manufacturer [8]. However, for the same type of analysis and detection items, the same analysis and detection principles are usually applicable, and the output original experimental data usually have similar compositions and structures. Therefore, this data filtering module can adapt and be applied to the original data generated by a flow analyser produced by another analytical device manufacturer by simply modifying the name and location of the analysis items.

Items	Maximum (mg/L)	LOD (mg/L)
<input type="checkbox"/> TN	17	0.04
<input type="checkbox"/> NO _x	7	0.015
<input type="checkbox"/> NO ₂	0.2	0.003
<input type="checkbox"/> NH ₄	12	0.046
<input type="checkbox"/> TP	1.2	0.023
<input type="checkbox"/> P04	1.2	0.023
<input type="checkbox"/> All N		
<input type="checkbox"/> All P		

Order No: ☐ Save separately

Start Repeat

Figure 2. Filtering module interface.

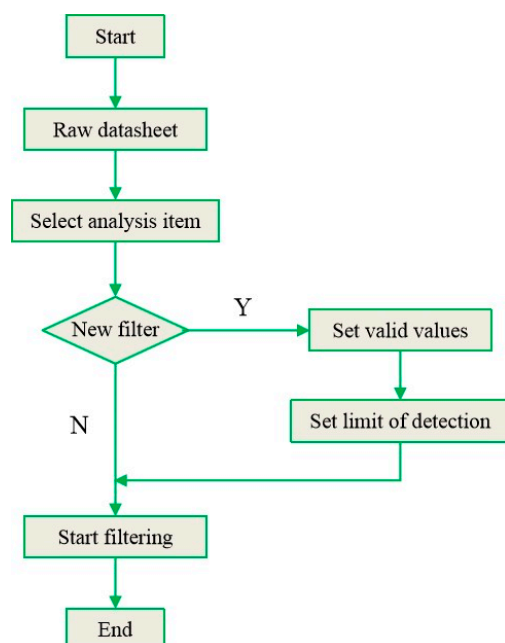


Figure 3. Data filtering workflow.

3.2. Data Quality Control

In the chemical analysis laboratory, instrument analysis plays an important role in the stable and accurate detection results of the instrument, from selecting the correct measurement principle to effectively monitoring each link of the measurement process. Therefore, analysts should strive to establish effective control and evaluation methods for the detection process, in order to prevent the occurrence of detection errors and ensure the reliability of the detection results [25].

In addition to the concentration value of the analysis item, the data file generated by the CFA after the experiment also contains a lot of other information, such as signal strength, peak time, and peak mark. For customers, this information is not required, and they do not need them in the data report; however, for analysers such information is usually necessary

to judge the stability and reliability of the state of the experimental instrument according to the data related to the detection signal. Therefore, the original data file needs to contain relevant information such as the original signal strength of the sample, peak time, baseline signal strength, quality control signal strength, sensitivity correction, etc. Therefore, we designed a quality control module through which analysers can capture the rich data of each analysis item and express the experimental quality visually. The analysers can evaluate the stability and reliability of the experiment through the standard curve, baseline state, sensitivity change, and other information. The quality of the experimental data can be evaluated through appropriate QC analysis. Feedback to the instrument analysis system can be used to adjust the parameters of the analysis system and improve the operating state and quality of the analysis system. The quality control module is shown in Figure 4.

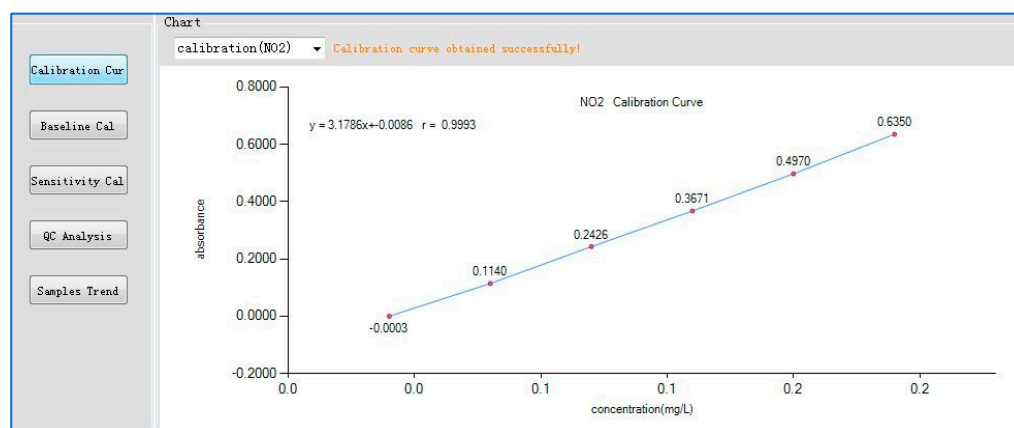


Figure 4. Data quality control module.

3.3. Data Report

After data filtering and processing, the data need to be further arranged into a standard format analysis report according to the specification requirements of the quality control system. A standard format data report is shown in Figure 5.

Analysis and Test Report		
Order No: YYYY-MM-DD-YY00xx		Report date: YYYY-MM-DD
		Page 1/Total 1
Sample number	Analysis item1 Unit1	Analysis item2 Unit2
1	xx. x	xx. x
2	xx. x	xx. x
3	xx. x	xx. x
4	x. xx	x. xx
5	x. xx	x. xx
6	x. xx	x. xx
7	xxx	xxx
8	xxx	xxx
9	xxx	xxx
10	xxxx	xxxx
Blank below		
Testing personnel: XXX Reviewer: XXX certifier: XXX		

Figure 5. Standard format data report structure. It has clear and specific requirements for row spacing, column spacing, and the number of rows per page. It also has clear requirements for the content, distribution, and size of elements, other than sample data.

When making a standard data report, analysts need to be very careful when filling the experimental data according to the standard template in an Excel spreadsheet. This is particularly important when there are many samples and many analysis items, as the spreadsheet will produce a cross-page distribution. In short, manually adjusting the data analysis report according to the standard data report format is very inconvenient and time consuming. Therefore, we have developed an automatic data report generation module to realise the automatic typesetting of the experimental data report and output the standard format data report that conforms to specifications. The module interface is shown in Figure 6.

Figure 6. Automatic data report generation module.

After entering the order number, tester, and approver into the module interface, we can manually set the number of rows and columns of each page below, or click automatic rows and columns so that the software module will automatically set the number of rows and columns of each page according to the numbers of rows and columns in the original data, and according to the requirements of the standard data report. Then, we click “generate report”, and the data report in standard format will be automatically generated and typeset. We click “print report” to automatically print the data report when a printer is connected.

3.4. Data Manage and Distribute

After analysing, processing, and compiling the experimental data to generate the standard report, the analysts have no unified management means and maintenance mechanism for their storage, maintenance, submission and distribution [26]. People save data report documents locally according to their personal habits, often without follow-up maintenance and disaster recovery measures, yet there can be data unavailability caused by unexpected hardware failure. The submission and distribution of the current document can only be carried out according to their respective information conditions, and the corresponding upload operation can only be carried out manually, one by one, according to the order number. When the number of documents is large, this process becomes time-consuming and laborious; there is no convenient method for the maintenance, archiving, and statistical analysis of historical documents. In other words, current data management lacks efficient submission and distribution means, nor are there means for low-latency, persistent, and highly reliable storage management and maintenance. In order to address this issue, we developed a data management module in which the client cooperates with the server to complete data management and distribution. The server uses Alibaba cloud ECS with low latency, durability, and high reliability, to achieve efficient reporting of data documents as

well as achieve reliable storage management and maintenance. The data report documents are archived locally and centrally, while the server maintains complete backups.

After the standard data report is generated, according to our business process, we need to distribute the reports to end customers and upload them to our LIMS for management. In the past, reporting data documents was done manually, one by one; this is cumbersome, time consuming and inefficient. According to the design of the LIMS system module, analysts must manually click on the corresponding position of the analysis and test business list page in order to open the analysis result page of a specific business, then select the corresponding data document, save and submit, and complete the submission. The operation steps of an order can require more than ten steps. When there are many orders, this way of manually submitting data documents one by one is cumbersome, time consuming, and inefficient. Since the upload task is executed by the local computer in real time, it will occupy the local computer and bandwidth resources of the analyst. In order to address this issue, we used an automatic uploading method for customer data in the LIMS system under the user mechanism in the module design [27], which is based on the Client/Server (C/S) architecture design, and is designed by Microsoft Visual Studio with its integrated development environment. The workflow is shown in Figure 7.

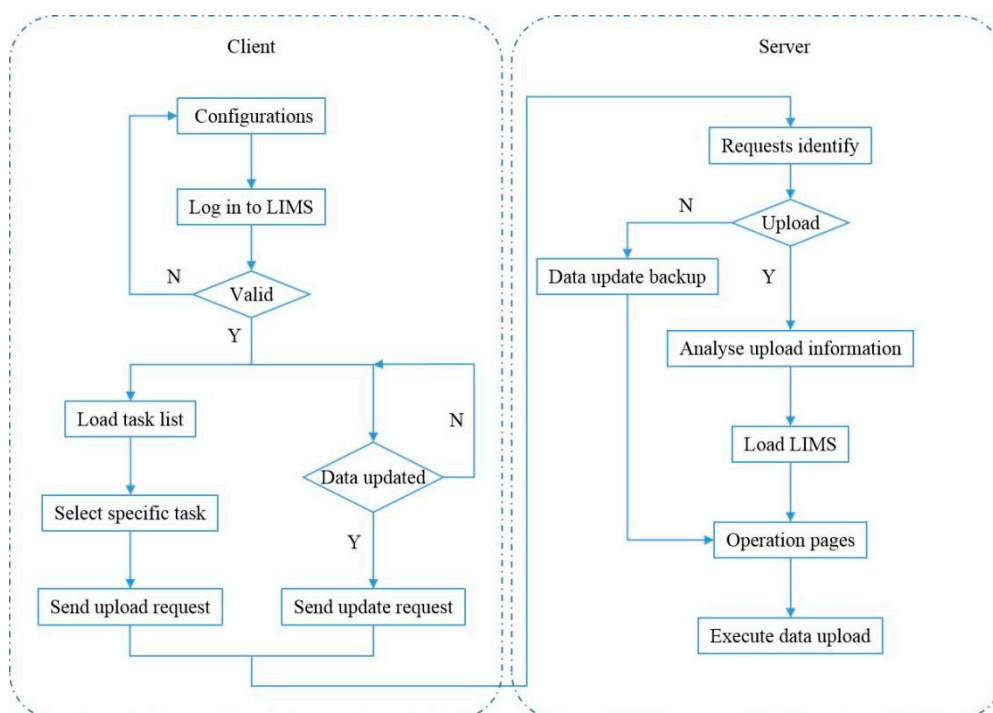


Figure 7. Workflow of an automatic uploading method for customer data in the LIMS system under the user mechanism.

The client realizes login and loads the corresponding business module of the LIMS system. When the local data document is updated, it will be automatically backed up to the server (S) through the client (C), and the data document upload request of a specific business will be submitted to the server. The data backup and upload requests are executed in parallel with multiple threads, and the local computer and broadband resources are reasonably allocated to improve the execution efficiency of the client. The server receives the data document backup request from the client and cooperates to realize the data document backup, receive the client-side data document upload request, and select the corresponding business data document on the server side according to the request information to automatically upload to the corresponding business module of the LIMS. By identifying and locating the page button position of LIMS-related business modules, the server simulates the implementation of the manual mouse click operation, completes

the upload of data documents, and automates data uploading. The server system has rich resources and bandwidth, which can quickly realize and complete the upload operation and data communication with the LIMS system. It can respond to multiple client requests at the same time. The experimenters only need to send a few request command bytes to the server at the client, and the server will automatically complete data report processing without occupying their local computer and bandwidth resources. The software module interface is shown in Figure 8.

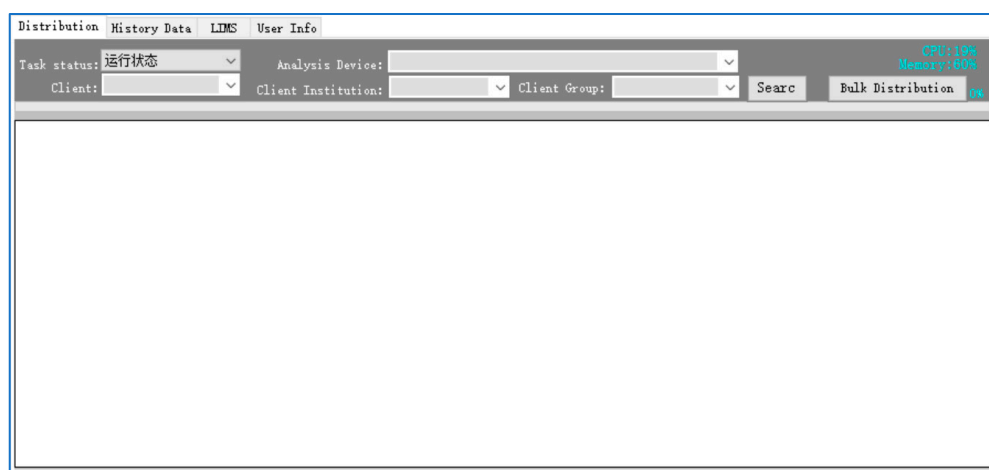


Figure 8. Data Manage and Distribute module interface.

4. Results and Discussion

The flexible data evaluation system designed in this paper was developed according to modular functions. Each module can run independently to fulfill its function, and can also be integrated for application. The integrated system is shown in Figure 9.

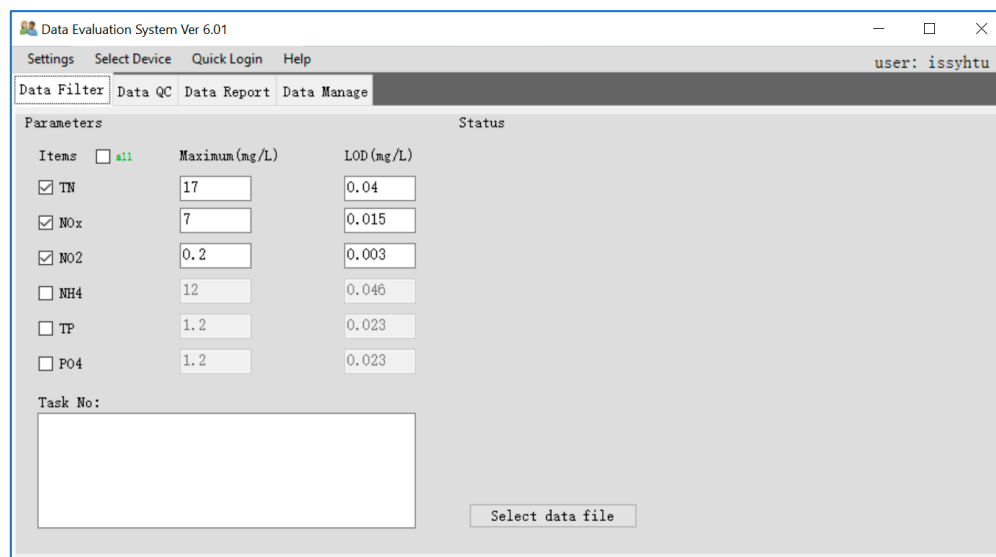


Figure 9. Flexible data evaluation system.

In the flexible data evaluation system, the processing targets of the data filtering module and the data quality control module are the original data derived from the instrument analysis system. The processing results of the data quality control module are used by analysts to evaluate the stability and reliability of the experiment and are fed back to the instrument analysis system to adjust the parameters of the analysis system and improve the operation state and quality of the analysis system. The processing object of the stan-

standard report module is the data document output after data filtering, and finally, the data report module outputs the standardised standard format data report, which is the final analysis result required by customers. The processing object of the data management and distribution module is the standard data report, which realizes storage management and maintenance with low delay, durability, and high reliability; moreover, it cooperates with the existing LIMS system to complete the efficient distribution of data reports to customers.

In this section, we verify the application and effect of each module in the business according to the business process of using the CFA for detection. Following analyses, we export all the original data from the device as the source data for the next step of processing; we select Excel Sheet 1 containing all the original data, run the software system, and check the analysis items for TN, NO_x, NO₂, and NH₄, in turn. The effective value is the maximum value effectively measured under the initial weight of the device, which is set according to the maximum value of the standard curve and the analysis experience of the device. The LOD is the minimum value effectively determined by the CFA. After selecting the analysis items, the start processing button is clicked, and the software will automatically filter, identify, and mark the source data. The processed results are placed in the new sheet. The results after data filtering are shown in Figure 10.

Sample No.	TN	Nox	NO ₂	NH ₄	QC data	Sample	TN-Results	NOx-Results	NO ₂ -Results	NH ₄ -Results
	mg/L	mg/L	mg/L	mg/L	Position	Identity	[mg/L]	[mg/L]	[mg/L]	[mg/L]
1 (2021-06-04-YY0002)	1.43	0.714	0.010	0.173	ST1	Drift	12.26132	4.47485	0.16015	7.97259
2	1.56	0.848	0.014	0.196	WT	Wash	0	-0.08963	0	0.03469
3	1.36	0.760	0.015	0.143	ST1	Drift	12.10004	4.97599	0.16176	7.99914
4	1.38	0.670	0.010	0.147	WT	Wash	0	-0.08963	0	0.03469
5	1.34	0.599	0.007	0.180	ST1	Drift	13.58883	4.45431	0.16051	7.70453
6	1.45	0.687	0.021	0.193	WT	Wash	0	-0.08963	0	0.03469
7	1.42	0.646	0.010	0.119	ST1	QC	12.24471	4.13747	0.1605	7.99242
8	0.909	0.399	0.009	0.165	ST1	Drift	12.16738	4.50888	0.16209	7.58583
9	1.22	0.616	0.008	0.104	WT	Wash	0	-0.08963	0	0.03469
10	1.45	0.632	0.015	0.101	ST1	Drift	14.3534	4.45814	0.16378	7.51351
11	1.22	0.598	0.011	0.0697	WT	Wash	0	-0.08963	0	0.03469
12	1.77	0.780	0.016	0.147	ST1	QC	12.54535	4.08437	0.16058	7.85292
13	3.22	2.12	0.003	0.100	ST1	Drift	12.02361	4.48119	0.16299	7.37311
14	3.83	2.36	0.005	0.121	WT	Wash	0	-0.08963	0	0.03469
15	4.40	2.79	0.005	0.0889	ST1	Drift	14.37448	4.57792	0.16429	7.35489
16	3.67	2.24	0.005	0.130	WT	Wash	0	-0.08963	0	0.03469
17	5.35	3.55	0.011	0.128	ST1	QC	12.44196	4.06776	0.16072	7.95157
18	5.70	3.71	0.008	0.205	ST1	Drift	12.28945	4.6297	0.16353	7.19982
19	4.72	2.87	0.008	0.151	WT	Wash	0	-0.08963	0	0.03469
20	4.79	3.04	0.009	0.116	ST1	Drift	14.4291	4.51523	0.16381	6.99801
21	4.37	2.65	0.008	0.197	WT	Wash	0	-0.08963	0	0.03469
22	4.31	2.51	0.012	0.168	ST1	QC	12.27895	4.03659	0.1606	8.02426
23	4.44	2.60	0.013	0.184	ST1	Drift	12.09638	4.5978	0.16396	7.05014
24	4.43	2.64	0.012	0.144	WT	Wash	0	-0.08963	0	0.03469
25	5.36	2.98	0.025	0.0999	ST1	Drift	13.10369	4.48584	0.18236	6.76856
26	5.11	2.78	0.023	0.172	WT	Wash	0	-0.08963	0	0.03469
27	5.11	2.81	0.026	0.124	ST1	Drift	12.59533	4.48201	0.17012	6.98652
28	2.56	1.04	<0.003	0.373	WT	Wash	0	-0.08963	0	0.03469
29	2.35	1.03	<0.003	0.351	ST1	QC	12.52345	4.05735	0.16054	7.83472
30	1.74	0.767	<0.003	0.328	ST1	Drift	12.59075	4.51846	0.16679	6.39813
31	2.29	1.19	<0.003	0.207	WT	Wash	0	-0.08963	0	0.03469
32	2.35	1.11	<0.003	0.235	List of samples exceeding valid value					
33	2.42	1.08	<0.003	0.341	D41	1 (2021-06-04-YY0002)	78.8	0.0237	0.020	34.0
34	2.55	1.15	<0.003	0.371	D42	2	47.0	0.0484	0.026	34.0
35	2.79	1.29	<0.003	0.350	D43	3	128	0.128	0.032	33.9
36	2.83	1.35	<0.003	0.340	D44	4	35.6	0.0471	0.031	33.9
37	3.95	2.75	0.007	0.174	D45	5	121	0.0463	0.019	33.9

Figure 10. Data processing results. After processing the raw data, all samples are placed in sequence, and the values are processed according to three significant digits. The quality control data and the samples that exceed the standard to be diluted are displayed separately, and all processes are completed automatically. Cells with a yellow background indicate that the value exceeds the maximum effective value of the analysis item and needs to be analyzed again after dilution. Cells with a red background indicate the number of the sample and the position information on the sample disk of the CFA.

The automatic processing time for 220 sample data in this experiment is 50 s; this usually takes at least 10 min for manual processing. Therefore, after using automatic data filtering, the data processing speed is greatly improved, and labour costs are saved. Along

with the reliability of machines and equipment being constantly improved, the focus of system reliability has switched away from human reliability analysis. Most accidents are caused by human errors [2]. Therefore, we can also avoid human errors by using automatic data filtering.

Similarly, Sheet1 containing all of the original data is selected, the data quality control module is run, and the experimental quality and data quality are visually expressed through the rich data of each analysis item. This output result is for analysts, as shown in Figures 11–14. The stability and reliability of the experiment can be evaluated through information on the standard curve, baseline state, and sensitivity change. The quality of the experimental data can be evaluated through an appropriate QC analysis. Based on this, analysts can adjust the operating parameters of the instrument analysis system in order to improve its operating state and quality.

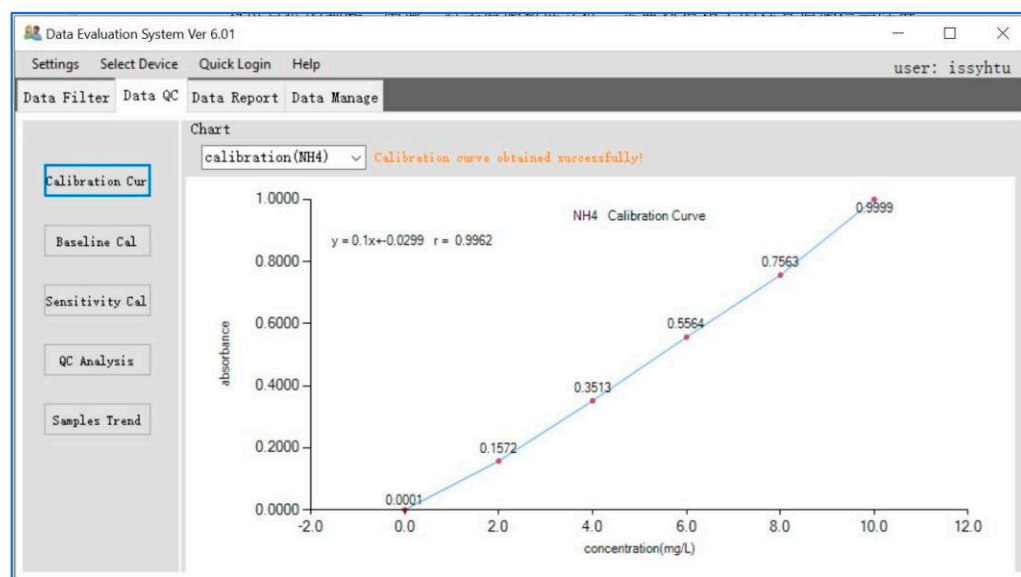


Figure 11. Detection calibration curve.

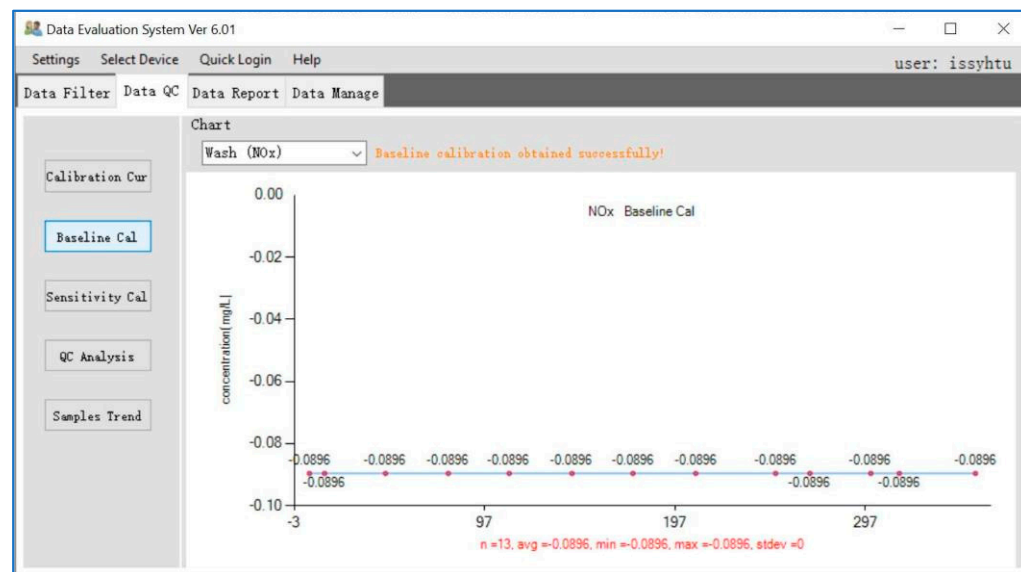


Figure 12. Baseline calibration of the sample data.

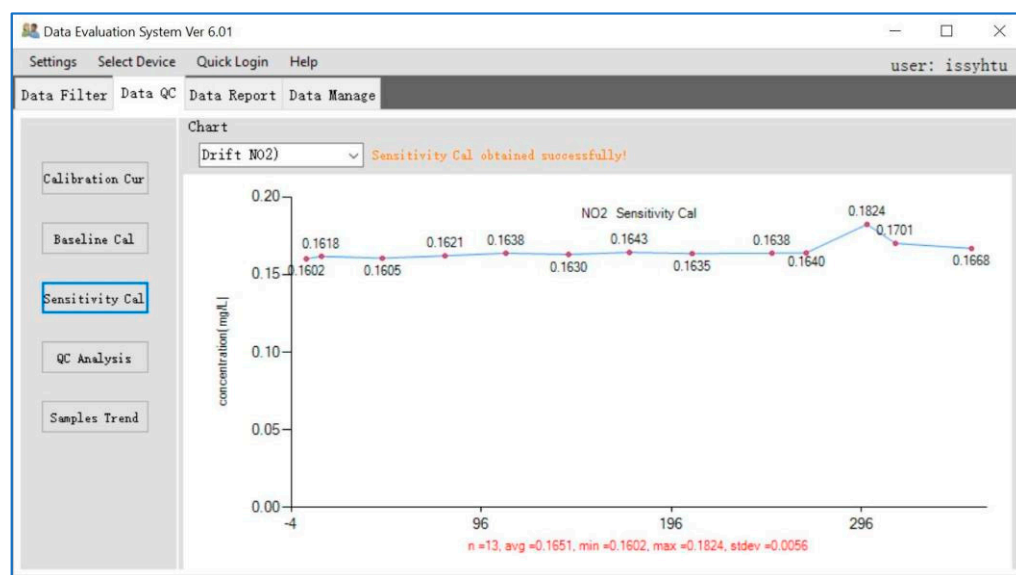


Figure 13. Sensitivity calibration of the sample data.

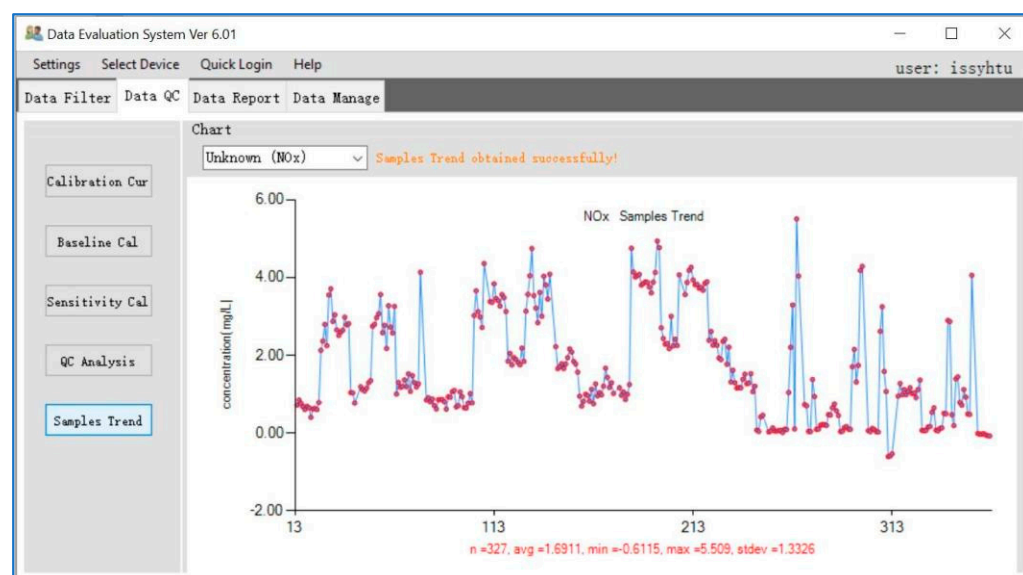


Figure 14. Sample data trend of the detection.

The filtered and processed data are stored in different sheets according to the business order number. According to the business process, it needs to be further arranged into a standard report format according to the specification requirements of the quality system. Double click the blank of the input column in the data report module to automatically fill in the order number, tester, and approver information; then, manually set the number of rows and columns of each page, or click automatic rows and columns for the software to automatically set the numbers of rows and columns of each page according to the numbers of rows and columns of the original data and the requirements of the standard report; click "generate report", and the standard format report as shown in the previous Figure 5 will be generated automatically. The standard report interface is shown in Figure 15. According to the number of samples and the number of analysis items, the module is recommended to be arranged into 30 rows and 2 columns

Analysis and Test Report									
Order No: 2021-06-04-YY0002 Report date: 2021-6-27 Page 1/Total 4									
Sample number	TN (Calculated by nitrogen)	NOx (Calculated by nitrogen)	NO2 (Calculated by nitrogen)	NH4 (Calculated by nitrogen)	Sample number	TN (Calculated by nitrogen)	NOx (Calculated by nitrogen)	NO2 (Calculated by nitrogen)	NH4 (Calculated by nitrogen)
	mg/L	mg/L	mg/L	mg/L		mg/L	mg/L	mg/L	mg/L
1	1.43	0.714	0.010	0.173	31	2.29	1.19	<0.003	0.207
2	1.56	0.848	0.014	0.196	32	2.35	1.11	<0.003	0.235
3	1.36	0.760	0.015	0.143	33	2.42	1.08	<0.003	0.341
4	1.38	0.670	0.010	0.147	34	2.55	1.15	<0.003	0.371
5	1.34	0.599	0.007	0.180	35	2.79	1.29	<0.003	0.350
6	1.45	0.687	0.021	0.193	36	2.83	1.35	<0.003	0.340
7	1.42	0.646	0.010	0.119	37	3.95	2.75	0.007	0.174
8	0.909	0.399	0.009	0.165	38	4.08	2.80	0.006	0.165
9	1.22	0.616	0.008	0.104	39	4.28	2.96	0.006	0.118
10	1.45	0.632	0.015	0.101	40	4.53	3.07	0.004	0.134
11	1.22	0.598	0.011	0.0697	41	5.17	3.56	0.005	0.125
12	1.77	0.780	0.016	0.147	42	3.76	2.58	0.004	0.150
13	3.22	2.12	0.003	0.100	43	4.12	2.77	0.004	0.173
14	3.83	2.36	0.005	0.121	44	3.23	2.17	0.003	0.0934
15	4.40	2.79	0.005	0.0889	45	4.85	3.27	0.005	0.146
16	3.67	2.24	0.005	0.130	46	4.06	2.72	0.005	0.0943
17	5.35	3.55	0.011	0.128	47	3.91	2.57	0.005	0.103
18	5.70	3.71	0.008	0.205	48	4.87	3.26	0.006	0.124
19	4.72	2.87	0.008	0.151	49	2.64	0.999	<0.003	0.462
20	4.79	3.04	0.009	0.116	50	2.97	1.30	<0.003	0.434
21	4.37	2.65	0.008	0.197	51	2.79	1.17	<0.003	0.376
22	4.31	2.51	0.012	0.168	52	2.55	1.20	<0.003	0.327
23	4.44	2.60	0.013	0.184	53	3.07	1.36	<0.003	0.374
24	4.43	2.64	0.012	0.144	54	2.73	1.19	<0.003	0.325
25	5.36	2.98	0.025	0.0999	55	3.19	1.52	<0.003	0.355
26	5.11	2.78	0.023	0.172	56	2.48	1.07	<0.003	0.320
27	5.11	2.81	0.026	0.124	57	3.13	1.47	<0.003	0.288
28	2.56	1.04	<0.003	0.373	58	2.92	1.30	<0.003	0.346
29	2.35	1.03	<0.003	0.351	59	2.68	1.18	<0.003	0.298
30	1.74	0.767	<0.003	0.328	60	2.79	1.26	<0.003	0.287
Testing personnel: YonghuiTu Reviewer: HaoyeTang certifier: WenyuHu									

Figure 15. Standard data report format according to the specification requirements of the quality system.

After the standard report is generated, it needs to be stored, maintained, and distributed. Previously, laboratory personnel kept the report documents locally according to their personal habits, which may have lacked follow-up maintenance and disaster recovery measures; there may have been data unavailability caused by unexpected hardware failure. With manual processing, the distribution of documents can only be manually uploaded to the LIMS, one by one, according to their respective information conditions and order numbers. In contrast, our data management and distribution module realises unified, highly reliable management means and maintenance mechanisms, as well as efficient automatic upload and distribution means. As shown in Figure 16, “Distribution” can query the current business status of the analyst, including approved, sample registered, and analysed results. In order to do this, we initially check the order whose status is sample registered. After configuring parameters such as the LIMS account, “Bulk Distribution” is clicked to automatically match multiple data standard reports stored in the specified path to the

corresponding order for uploading and submission to the LIMS. This avoids the troubles involved with manually uploading report documents to the browser [27].

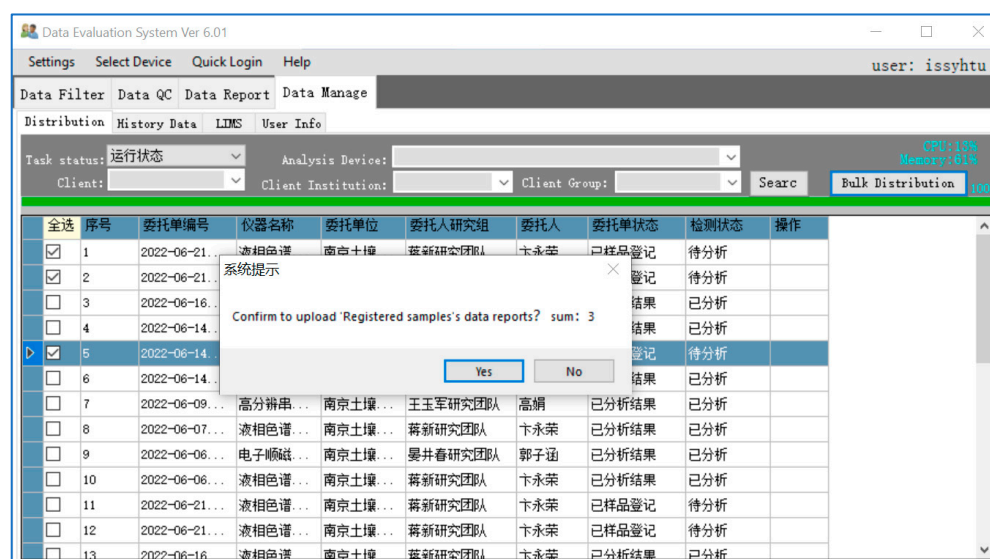


Figure 16. Data Manage and Distribute module.

5. Conclusions

The flexible data evaluation system designed in this paper was developed according to modular function. The functions of these four software modules are independent and can be implemented separately or in combination. According to the selected analysis items and the set maximum value and LOD, the data filtering module filters out data that are irrelevant to the customer from among a large number of original data, and automatically and quickly filters out the data required by the customer. From the test results, compared with the traditional manual screening method, the speed of this process was improved to be dozens of times faster, with errors introduced by human handling avoided. The processing object of the data report module is the data document output after data filtering. Through automatic typesetting and compilation, the final automatic output of the standard format data report meets all specifications for the final analysis result required by customers for sample instrument analysis. Compared with manual typesetting, it is efficient and unified without deviation. The data management and distribution module realize a unified management means and maintenance mechanism. It cooperates with the existing LIMS to upload and distribute data reports. The automatic uploading method of customer data in the LIMS system under the user mechanism is adopted, which overcomes the manual approach of submitting data documents one by one. This method is cumbersome and inefficient, and moreover occupies local computer and bandwidth resources when uploading a large amount of data is involved.

In addition, the modular design idea proposed in this paper enables the flexible data evaluation system to be easily extended and applied to different analysis systems produced by different analysis device manufacturers. By simply modifying the name and location identification of the analysis items, the data filtering module developed in this paper can be adapted and applied to the original data generated by the CFA produced by other analysis device manufacturers. The processing object of the data report module is the filtered data, and the processing object of the data management and distribution module is the standard data report; they have no direct correlation with the analysis system of the device and the output raw data. Therefore, even the output data evaluation of different types of device analysis systems can be integrated into the data evaluation system merely by developing the data filtering module through separate adaptation. Thus, the application scope of the data evaluation system can be expanded. In fact, in subsequent research, we

also developed data screening modules that were suitable for different analysis systems, such as HPLC-ICP-MS, ICP-AES, and industrial nano CT, which were integrated into the data evaluation system to meet the needs of these instruments and equipment testing, and improve the analysis and testing quality and efficiency.

In summary, this paper combed and analysed the data evaluation work in the analysis and testing business of traditional chemical analysis laboratories, and developed a flexible data evaluation system for CFA; realized automatic data filtering, quality control, data management and distribution; integrated with the existing LIMS to maximise automation; and improved the quality and efficiency of the analysis and testing business. The modular design idea makes it easy for the data evaluation system to be partially or entirely extended. It provides a reference for different analysis systems produced by different analysis device manufacturers, and provides a reference for improving the quality and efficiency of the analysis and testing business in chemical analysis laboratories.

6. Patents

The research involved in this paper secured a Chinese invention patent [27], which discloses an automatic uploading method of customer data in the LIMS system under the user mechanism, including the following steps: S1: the client stores the customer data document in real time to the designated path of the local storage data document of the corresponding user; S2: the client regularly determines whether the document under the above path has been updated; if so, it sends a document update request to the server and transmits the updated document to the server, and sends a report request when data reporting is needed; S3: the server determines the received request information of the client; if it is an update request, it receives the update document sent by the client and saves it to the corresponding user path; if it is a submission request, the document is found under the corresponding path on the server side for submission. The invention only needs to select single or multiple experimental analysis businesses that need to report data documents in the client interface, and click the batch submit button for the client to send a report request to the server, and the data report business will be automatically completed by the server.

Author Contributions: Conceptualisation, Y.T. and H.T.; methodology, W.H.; software, Y.T.; validation, Y.T., H.T. and W.H.; formal analysis, H.G.; investigation, W.H.; resources, H.T.; data curation, Y.T.; writing—original draft preparation, Y.T.; writing—review and editing, W.H.; visualisation, H.T.; supervision, W.H.; project administration, Y.T.; funding acquisition, W.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (Grant number 2021YFD1500202), the Key-Area Research and Development Program of Guangdong Province, China (Grant number 2020B0202010006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Marada, W.; Poniszewska-Marada, A.; Szymczyńska, M. Data Processing in Cloud Computing Model on the Example of Salesforce Cloud. *Information* **2022**, *13*, 85. [[CrossRef](#)]
2. Zhao, C.; Ding, Y.; Yang, Z. Studies on human error and its identification technique. *Ind. Saf. Environ. Prot.* **2002**, *5*, 40–43.
3. Bao, Y.K.; Wang, Y.F.; Wen, Y.F. Equipment Reliability Evaluation and Maintenance Period Decision Considering the Impact of Human Factors. *Power Syst. Technol.* **2015**, *39*, 2546–2552.
4. Xu, X.; Zheng, W.; Gu, Y. Analysis and Correction of Error for colormeter System based on Neural Network. *Process Autom. Instrum.* **2006**, *2*, 25–27.
5. Jiang, Z.; Xu, M.; Yao, W. The Correction of Instrument System Errors based on Numerical Analysis. *Geophys. Geochem. Explor.* **2013**, *37*, 287–290.
6. Thompson, M. Precision in chemical analysis: A critical survey of uses and abuses. *Anal. Methods* **2012**, *4*, 1598–1611. [[CrossRef](#)]

7. Bosona, T. Urban Freight Last Mile Logistics-Challenges and Opportunities to Improve Sustainability: A Literature Review. *Sustainability* **2020**, *12*, 8769. [[CrossRef](#)]
8. Fleischer, H.; Adam, M.; Thurow, K. Flexible Software Solution for Rapid Manual and Automated Data Evaluation in ICP-MS. In Proceedings of the 32nd Annual IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Pisa, Italy, 11–14 May 2015; pp. 1602–1607.
9. Zhu, B.Y.; Wu, R.L.; Yu, X. Artificial Intelligence for Contemporary Chemistry Research. *Acta Chim. Sin.* **2020**, *78*, 1366–1382. [[CrossRef](#)]
10. Emerson, J.; Kara, B.; Glassey, J. Multivariate data analysis in cell gene therapy manufacturing. *Biotechnol. Adv.* **2020**, *45*, 107637. [[CrossRef](#)]
11. Qi, Y.L.; O'Connor, P.B. Data processing in Fourier transform ion cyclotron resonance mass spectrometry. *Mass Spectrom. Rev.* **2014**, *33*, 333–352. [[CrossRef](#)]
12. Urban, J.; Vanek, J.; Stys, D. Current State of HPLC-MS Data Processing and Analysis in Proteomics and Metabolomics. *Curr. Proteom.* **2012**, *9*, 80–93. [[CrossRef](#)]
13. Vervoort, Y.; Linares, A.G.; Roncoroni, M.; Liu, C.; Steensels, J.; Verstrepen, K.J. High-throughput system-wide engineering and screening for microbial biotechnology. *Curr. Opin. Biotechnol.* **2017**, *46*, 120–125. [[CrossRef](#)] [[PubMed](#)]
14. Leavell, M.D.; Singh, A.H.; Kaufmann-Malaga, B.B. High-throughput screening for improved microbial cell factories, perspective and promise. *Curr. Opin. Biotechnol.* **2020**, *62*, 22–28. [[CrossRef](#)] [[PubMed](#)]
15. Trapp, O. Gas chromatographic high-throughput screening techniques in catalysis. *J. Chromatogr. A* **2008**, *1184*, 160–190. [[CrossRef](#)] [[PubMed](#)]
16. Shukla, S.J.; Huang, R.L.; Austin, C.P.; Xia, M.H. The future of toxicity testing: A focus on in vitro methods using a quantitative high-throughput screening platform. *Drug Discov. Today* **2010**, *15*, 997–1007. [[CrossRef](#)]
17. Stephan, C.; Kohl, M.; Turewicz, M.; Podwojski, K.; Meyer, H.E.; Eisenacher, M. Using Laboratory Information Management Systems as central part of a proteomics data workflow. *Proteomics* **2010**, *10*, 1230–1249. [[CrossRef](#)]
18. Cao, C. Different Analysis of Instrumental Analysis and Chemical Analysis. *Chem. Eng. Des. Commun.* **2016**, *42*, 43.
19. Sinenian, N.; Zylstra, A.B.; Manuel, M.J.E.; Frenje, J.A.; Kanojia, A.D.; Stillerman, J.; Petrasso, R.D. A Multithreaded Modular Software Toolkit for Control of Complex Experiments. *Comput. Sci. Eng.* **2013**, *15*, 66–75. [[CrossRef](#)]
20. Li, Z.; Cheng, L.; Yao, L.; Tong, X. Information Management of Data Resources in Product Design and Manufacture Process. In Proceedings of the International Conference on Manufacturing Engineering and Automation, Guangzhou, China, 7–9 December 2010; pp. 1674–1678.
21. Poniszewska-Maranda, A.; Grzywacz, M. Storage and Processing of Data and Software Outside the Company in Cloud Computing Model. In Proceedings of the 12th International Conference on Mobile Web and Intelligent Information Systems (MobiWis), Rome, Italy, 24–26 August 2015; pp. 170–181.
22. Ruzicka, J.; Hansen, E.H.; Ramsing, A.U. Flow-injection analyzer for students, teaching and research—Spectrophotometric methods. *Anal. Chim. Acta* **1982**, *134*, 55–71. [[CrossRef](#)]
23. Henning, W. Semiautomatic connective-tissue and total phosphate determination using the scalar continuous-flow analyzer. *Fleischwirtschaft* **1990**, *70*, 97–101.
24. Yang, J.; Zhang, Z.; Cao, G. Soil nitrate and nitrite content determined by Skalar SAN~(++). *Soil Fertil. Sci. China* **2014**, *2*, 101–105.
25. Sun, X.; Qi, X.; Hou, J. The quality control method of soil or fertilizer testing laboratory. *Chem. Anal. Meterage* **2022**, *31*, 78–83.
26. Wu, J.; Zuo, H.; Meng, L.; Yang, Y.; Cheng, Z.; Liu, H. IETM Data Management. In Proceedings of the Chinese Automation Congress (CAC), Hangzhou, China, 22–24 November 2019; pp. 5158–5162.
27. Tu, Y.; Tang, H.; Hu, W.; Sun, C. An Automatic Uploading Method of Customer Data in LIMS System under the User Mechanism. CN Patent CN109525642B, 15 September 2020.