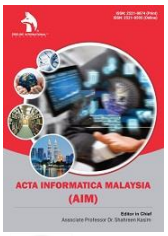
ZIBELINE INTERNATIONAL PUBLISHING

ISSN: 2521-0874 (Print)
ISSN: 2590-4043 (Online)
CODEN: AEMCDV

Acta Informatica Malaysia (AIM)

DOI: http://doi.org/10.26480/aim.02.2023.101.104



CrossMark

RESEARCH ARTICLE

DESIGN AND RESEARCH OF AUTOMATIC STATION BUFR DATA EMERGENCY BACKUP PLATFORM BASED ON HYBRID PROGRAMMING

Hao Liu[a], Jie Fu[a*], Peng Gao[a], Juan Xu[a]

[a]Gansu Provincial Meteorological Information and Technical Support and Equipment Center, Lanzhou 730000, China
**Corresponding Author Email: 420004048@qq.com*

ARTICLE DETAILS

Article History:

Received 21 March 2023
Revised 24 April 2023
Accepted 27 May 2023
Available online 03 June 2023

ABSTRACT

Relying on the domestic meteorological transmission and communication system and NAS backup mechanism, based on the relevant standards and specifications of the meteorological big data cloud platform, in the case that the "Tianqing" interface cannot obtain data in a timely and effective manner, detailed requirements, detailed research, and clear implementation steps are made to provide users with reliable meteorological data services. Through the design and research of automatic station BUFR data rain express emergency backup system in this paper, the BUFR data for cell turn decoding, the specific factor value processing, can provide users with continuous data service, for business personnel in flood season normal work to provide basic support. The application shows that the sustainability of the data service is enhanced, and it provides a strong guarantee for the acquisition of rainfall elements in flood season.

KEYWORDS

computer technology and applications; meteorological big data cloud platform; BUFR data; programming

1. INTRODUCTION

Meteorological data and information are the fundamental support for the development of the meteorological industry. Standardized format for meteorological data is beneficial for its transmission among various business systems worldwide (Chen et al., 2023). In 1947, the earliest standard format for weather code, the "International Meteorological Code", was developed by the Weather Committee under the promotion of the World Meteorological Organization and was put into use in 1949. This standard format adopted character encoding (TAC) that is suitable for the objective conditions for early data transfer, and has the advantages of intuitiveness and readability (Wang et al., 2022). With the change of business needs, the format needs continuous expansion and improvement as the encoding format is complicated, prone to repetition and errors (Li et al., 2022). Since 2000, with the rapid development of the Internet and computer technology, the World Meteorological Organization has developed the table-driven code format (TDCF) based on binary and vigorously promoted it to adapt to the development needs. TDCF includes commonly used BUFR, GRIB, and CREX codes (Zhao et al., 2022). Between 2003 and 2014, the encoding conversion and transitional work for various meteorological data were gradually completed. In 2014, under the leadership of the China Meteorological Administration, China deployed the standardization work of meteorological data formats (Chen et al., 2022). By 2020, China has basically achieved the goal of standardizing meteorological data formats. Currently, to better meet the actual needs of the domestic meteorological industry, China's meteorological data formats adopt a combination of internationally common data formats and domestic customized data formats (Jiang et al., 2023) and further optimize the technical route of standardizing meteorological data formats.

Currently, the meteorological industry in China uses the domestic meteorological transmission and communication system (CTS) for the transmission of meteorological data. For ground data, the transmission and archiving format from the production end (various meteorological stations) to the usage end in various systems has been converted to the BUFR code, which ensures the standardization and uniformity of data formats. BUFR format data transmission method has great practicality in the background of fully considering the complexity and specificity of meteorological service business, and maximizes the real-time and reliability of meteorological data (Zhang, 2010). With the support of the Meteorological Big Data Cloud Platform ("Tianqing"), the whole business process is realized, including observation data generation system - domestic meteorological transmission and communication system - quality control decoding system - data processing system - storage service system - Protal application system. The advantages of the integration of meteorological data formats are more obvious. Relying on Tianqing, using the interface application development based on MUSIC-related standards and specifications, rich and reliable meteorological data services are provided to users at the usage end (Zhang & Yin, 2021).

As a business personnel in the meteorological industry, the specific situation in Gansu province presents challenges during the flood season in terms of providing effective support. As the number of visits to the "Tianqing" interface increases sharply, it often leads to blockages and impediments in interface access (Gao et al., 2018). Moreover, with regards to data loading, any impact on the XuGu database may result in disabled interface access to the necessary information. In the first scenario, users may experience a delay in accessing relevant information, rendering it impossible to provide a prompt and efficient data service. In the latter case, users and business personnel are forced to deal with a lack of available data. After obtaining ground data in BUFR format from the CTS or NAS backup system, it is still impossible to directly retrieve the corresponding numeric values of relevant parameters from the message. Consequently, business personnel urgently need a platform that can promptly provide the actual numeric values of relevant parameters from the BUFR format of ground data, even during database or interface issues. The proposed emergency backup platform for automatic station BUFR data offers an effective solution to this issue. By adding an additional layer

Quick Response Code	Access this article online	
	Website: www.actainformaticamalaysia.com	**DOI:** 10.26480/aim.02.2023.101.104

Cite the Article: Hao Liu, Jie Fu, Peng Gao, Juan Xu (2023). Design and Research of Automatic Station Bufr Data Emergency Backup Platform Based on Hybrid Programming. *Acta Informatica Malaysia, 7(2): 101-104.*

of support to the existing operational workflow, this platform ensures the continuous provision of relevant data services to business personnel even when other methods of information retrieval have failed. This greatly alleviates business personnel's work pressure, and its implementation holds positive implications for the meteorological industry.

2. Demand Analysis

The "Tianqing" system provides "cloud+ terminal" business support capabilities for both national and provincial level, and solves the data support environment for business personnel under the intensive standardization. However, due to the following reasons, there are still challenges to be addressed: Firstly, the rapid increase in the number of visits to the "Tianqing" interface leads to blockages and impediments in interface access. Secondly, any impact on the XuGu database in terms of data loading may result in disabled interface access to the necessary information. Under these circumstances, business personnel and scientific research users may encounter difficulty in accessing data due to their lack of familiarity with data inventory, observation elements and storage formats.

Regarding the specific situation in Gansu province during the flood season, business personnel places great importance on obtaining precipitation data. Currently, ground data is transmitted in BUFR format, including hourly and minute data from both automatic and regional stations. This BUFR-format data is transferred within the system and backed up in the NAS system. However, in the event of interface failure in the "Tianqing" system, users and organizations would be unable to efficiently and promptly obtain precipitation element values from the BUFR data. This situation must be addressed to ensure that necessary data can always be accessed to support effective flood season preparations and operations in the province.

When processing data, scientific researchers with lower business development skills may face challenges in obtaining data information in a timely and effective manner. Under such circumstances, they would typically need to explore the characteristics of the data format and look for corresponding decoding tools to perform secondary development, which increases both development and data usage costs (Dong et al., 2022). At the same time, business or scientific research personnel generally have a limited understanding of the structured and unstructured storage rules of the data, which indirectly makes it difficult for them to obtain data in the most effective way possible(Liu et al., 2021; Gao et al., 2022).

To address the issue of scientific researchers with different levels of business development skills unable to obtain data information in a timely and effective manner, the following measures can be taken: 1. Provide standardized data formats and data interfaces to facilitate access to the required data by scientific researchers. 2. Provide tools and APIs for parsing data to reduce the cost of data usage. 3. Establish a data sharing platform to facilitate data exchange and sharing among different researchers. 4. Introduce professional data processing tools and artificial intelligence algorithms to improve work efficiency and the value of data utilization. By adopting these measures, data resources can be more effectively utilized, promoting progress in scientific research and business development (Xu et al., 2022).

3. Technical Route

In practical operations, the complete workflow of information based on the "Tianqing" system is as follows: Firstly, meteorological data is collected from various stations or satellite broadcast systems. This data then undergoes rapid quality control through the domestic meteorological transmission communication system. Following this, the data is subjected to distribution processing, facilitating the dissemination of raw information to the various user systems (Liu et al., 2023; Zong & Zeren, 2019) In parallel, the data flows towards the downstream system, where it undergoes storage processing and decoding entry via a state-of-the-art decoding and storage management system. It is noteworthy that the storage management and product processing systems operate simultaneously, sharing an input and output interface. Thus, they not only engage in data treatment but also administer the information storage structure and strategy. Ultimately, through its robust interface, the system provides unparalleled data services to users, ensuring efficient and effective data management. The precise workflow of "Tianqing" is illustrated in Figure 1.

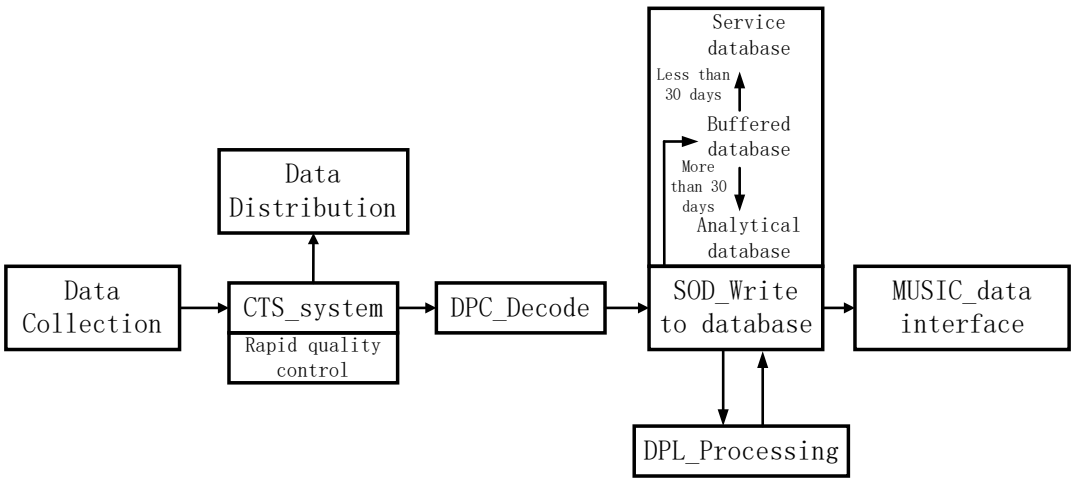


Figure 1: Schematic diagram of the existing simplified workflow of "Tianqing"

Based on the aforementioned discussion and practical experience, it is evident that there are three significant factors that impact user experience and usage in the data collection process from the "Tianqing" interface. Firstly, malfunctions and failures in the XuGu database can significantly impede the data loading process, making it impossible for users to obtain data through the interface. Secondly, extremely high interface traffic can cause bottlenecks, affecting the timeliness and availability of data access and collection or even making it entirely unavailable. Finally, errors in the data decoding and loading system can lead to decoding or data loading errors, further degrading user experience. All three scenarios can adversely impact user experience and, in severe cases, even affect the operations of several organizations. Furthermore, the workflow diagram indicates that the overall system follows a linear, single-flow structure from data collection to endpoint user interface, without a redundant emergency backup system to ensure data reliability. Therefore, this paper proposes a straightforward emergency backup system to supplement the existing workflow and enhance data reliability. The proposed backup system's functional workflow is depicted in Figure 2.

With reference to the proposed workflow diagram, the primary research focus is on the precipitation element value in BUFR data. There are two main ways of acquiring data: first, the automatic station hourly BUFR data is retrieved and decoded from the NAS backup disk, and second, data is obtained from the "Tianqing" interface. To avoid the time-consuming and risky process of obtaining data from external interfaces, it is recommended to prioritize the use of locally backed-up data. To this end, the NAS backup mechanism can be leveraged to understand user needs, become familiar with protocols such as ftp and sftp, and efficiently retrieve BUFR-format data from the NAS disk. The data processing mainly involves the use of the Pandas tool in Python to accumulate data for all stations for any desired time period. The cumulative values are then sorted in descending order. Finally, data such as the station number, station name, data time, precipitation, and accumulated precipitation are collected and presented in a tabular format, which can serve as a convenient reference for business personnel in generating precipitation reports.

Before the proposed emergency backup system can be fully integrated into practical applications, it is crucial to conduct a comprehensive data comparison and validation process. This involves verifying the accuracy of processed precipitation data generated by the backup system, using

known and reliable data sources such as meteorological stations and satellite data for comparison (Ren, 2020). By comparing the processed precipitation data with these reliable data sources, the accuracy of the backup system's processing and data generation can be evaluated. Any discrepancies or inconsistencies can be identified and addressed, leading to further refinement and improvement of the emergency backup system's data processing reliability and accuracy.

The emergency backup platform for automatic station BUFR data, based on hybrid programming, utilizes a C/S architecture and primarily serves meteorological business personnel, with a high demand for real-time data processing. Given the large number of automatic and regional stations in the province, the platform must process significant amounts of data through grid conversion and decoding. To enhance efficiency, the platform employs a multi-threaded grid conversion strategy with file locks and a parallel asynchronous processing approach using a thread pool for data processing. In this heterogeneous, non-clustered system, hybrid programming in Python and shell scripting is used to adapt to user data processing needs, ensuring data service continuity and reliability. The task decomposition steps for the platform design are shown in Figure 3.

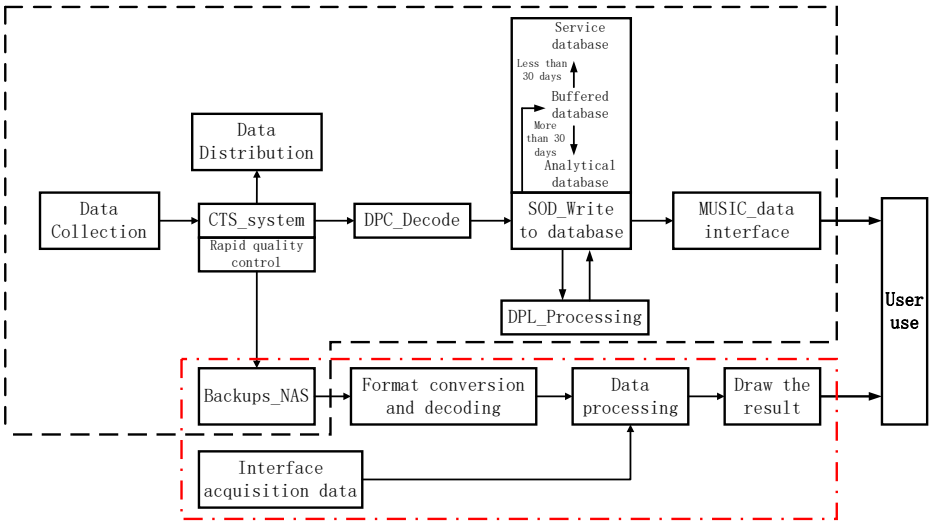


Figure 2: Schematic diagram of workflow of the emergency backup system

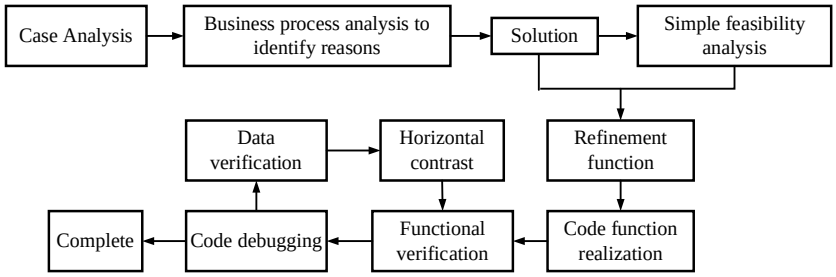


Figure 3: Steps of platform design task decomposition

4. Business Application

4.1 Performance Test and Comparison Verification

Indeed, in practical meteorological applications, simply providing raw meteorological data is not enough to provide accurate and reliable services. Rather, it is necessary to process and analyze the data utilizing statistical methods and applying meteorological expertise to understand the significance of the values. For example, in the case of precipitation data, it is common to categorize precipitation events using a set of rules. These rules may define events such as rainfall exceeding 50-99.9 millimeters over a period of 24 hours as "heavy rain", rainfall between 100-249.9 millimeters as "severe rain", and rainfall above 250 millimeters as "extreme rain". Using such categorization methods and other statistical tools can enhance our understanding and assessment of meteorological data, providing a more accurate basis for decision-making. Through simple statistical analyses and the application of meteorological knowledge and experience, it is possible to analyze the relationships between various meteorological factors and provide valuable insights for decision-making.

Meteorological personnel primarily use 5-minute and 10-minute precipitation accumulation values when analyzing precipitation data. To compare the actual precipitation values, precipitation data was obtained on July 24 and July 28 from the Lanzhou regional station, the Linxia regional station, and the automatic station in Gansu Province. Performance testing and data comparison were then conducted using the emergency backup platform for automatic station BUFR data based on hybrid programming. The results of the performance testing are shown in Table 1, and the comparison verification results are shown in Figure 4.

Table 1: Test Results of Time Consuming Performance of Precipitation Element Calculation

Data type	Number of files	Time consumed/min
Precipitation accumulation in 10 minutes at Lanzhou meteorological regional station on July 24	16380	1.14
Precipitation accumulation in 5 minutes at Lanzhou meteorological regional station on July 24	16380	1.23
Precipitation accumulation in 10 minutes at Linxia meteorological regional station on July 24	17712	1.15
Precipitation accumulation in 5 minutes at Linxia meteorological regional station on July 24	17712	1.29
Precipitation accumulation in 10 minutes at Gansu meteorological automatic station on July 24	116640	1.68
Precipitation accumulation in 5 minutes at Gansu meteorological automatic station on July 24	116640	1.73

According to the results shown in Table 1, when testing the performance of minute-level precipitation data with a daily time series, the processing time was all within 2 minutes due to the large amount of data. However, the emergency backup platform for automatic station BUFR data based on hybrid programming ran stably without any abnormal failures. In actual work scenarios, meteorological personnel typically require time series that are less than one day. With a control time of any duration less than 0.5 days, the processing time of the platform can be controlled within 1 minute, which satisfies the normal use requirements of the meteorological personnel.

The data sources used for the consistency comparison include the 10-

minute precipitation accumulation values obtained from the integrated platform and the 5-minute and 10-minute precipitation accumulation values calculated from the "Tianqing" interface precipitation data. As shown in Figure 4, the consistency of the precipitation data calculated from regional stations is poor. After investigating the platform, it was found that missing data records in the original data were not processed, resulting in some accumulated data processing results that did not match the actual values. After adding the function of rapid quality control processing for raw data on the platform, the consistency of the data was significantly improved.

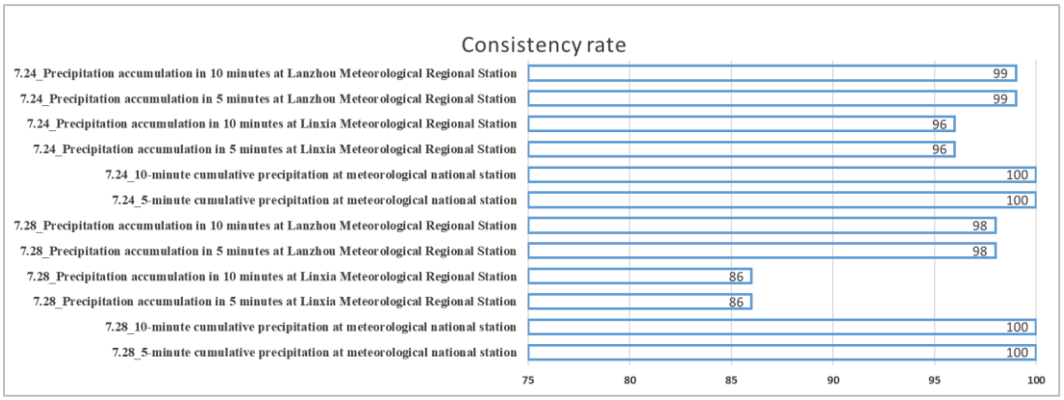


Figure 4: Comparison and verification results of data consistency

4.2 Typical Application

The emergency backup platform for automatic station BUFR data based on hybrid programming can provide services for obtaining specific element data, statistical analysis, and visualization of precipitation data. The platform can effectively obtain regional station and automatic station BUFR format data from the NAS storage system and process it to obtain the corresponding precipitation element values. The feasibility of the design process has been verified in practical applications, and it has also played a promoting role in the promotion and application of data interface services. Taking the precipitation data of Gansu Province on July 24, 2022 as an example. In addition, rainfall data can be dynamically monitored according to the time period.

5. Conclusions

The application of meteorological information in various aspects necessitates robust data support. In particular, during the flood season, timely access to reliable rainfall data is crucial for making informed judgments about the current situation. Failure to obtain such data expeditiously could have deleterious effects on many operational activities. Although data backup mechanisms are in place, much of the data transmission employs BUFR format, which could hinder obtaining veritable meteorological element values in emergent situations and thereby impede further operations.

The application of an automatic station BUFR data emergency backup platform, based on hybrid programming, has effectively alleviated the aforementioned issues. This development has paved the way for the progressive rollout of corresponding element extensions, which further enhances the functionality of the emergency backup system. This approach improves the operational stability of flood season business activities, ensures data accuracy and timeliness, optimizes business process flows, enhances work efficiency, and improves overall data service quality.

References

Chen, X.H., Liu, Y., Xu, J., et al., 2023. Research and application of multi-server distributed download technology based on FTP protocol. Electronic Design Engineering, 31(7), 42-46.

Chen, X.H., Liu, Y., Zhang, H., et al., 2022. Research and implementation of application management model in Linux large-scale cluster. Electronic Design Engineering, 30(10), 17-21, 26.

Dong, L.M., Li, Y.Z., Qin, Y.F., et al., 2022. Development and Interface Application Skills of Meteorological Service Client System Based on CMADaaS Platform. Meteorological Science and Technology, 50(2), 297-302.

Gao, P., Liu, Y., Fu, J., et al., 2022. Design and research of WebGIS platform for meteorological big data analysis based on Meteoinfo. Electronic Design Engineering, 30(15), 176-180.

Gao, Y.J., Liu, Y.C., Zhang, L., et al., 2018. Design and implementation of high-altitude detection data transmission software based on C#. Journal of Meteorological Research and Application, 39(2), 90-93.

Jiang, S.C., Zhuo, X.Y., Xie, Y.H., et al., 2023. Design and implementation of automatic filling and adjustment delivery program for ground meteorological station BUFR data. Straits Science, (1), 71-74, 93.

Li, G.J., Guo, J., Qiao, W.W., et al., 2022. Design of meteorological data download platform based on Tianqing system. Guangdong Meteorology, 44(6), 68-72.

Liu, X., Jia, L.D., Chen, X.J., et al., 2021. Introduction to the service interface system of meteorological big data cloud platform "Tianqing". Shanxi Meteorological Quarterly, (3), 105-108.

Liu, Y., Huang, Z., Xu, J., et al., 2023. Monitoring and Warning System of Meteorological Big Data Cloud Platform. Computer Systems & Applications, 32(3), 86-94.

Ren, W.J., 2020. Design of digital media data storage and sharing system based on MVC mode. Electronic Design Engineering, 28(4), 118-122.

Wang, X.D., Xu, J., Zhang, H., 2022. Implementation method for non-programming user access to meteorological data interface based on Tianqing. Gansu Science and Technology, 38(10), 18-21.

Xu, J., Liu, Y., Wang, X.D., et al., 2022. Implementation and application of backward compatible version meteorological service interface system. Science and Technology & Innovation, (18), 152-153.

Zhang, E.H., & Yin, H.Y., 2021. Decoding and analysis of meteorological satellite data based on compressed format BUFR code. Guangdong Meteorology, 43(5), 70-74.

Zhang, X., 2010. A new method for converting meteorological BUFR data model to GIS data model. In: Proceedings of the Fifth Academic Forum for Geography Graduate Students of Beijing Area. 30-31.

Zhao, Y.J., Jiang, H.S., Li, Z.F., et al., 2022. Research of Intelligent Grid Forecast Data Sharing Method Based on Sky Engine. Tianjin Science and Technology, 49(11), 49-51.

Zong, Z, Zeren Dawa., 2019. Operation problem and treatment of new ground comprehensive observation service software of automatic weather station. Agriculture and Technology, 39(13), 146-147.