Provincial Data Quality Control Strategy for the Operating Platform of Meteorological Data Business System

Xuanyu Zhang*, Bingyu Zhao, Xinhua Huang

Meteorological Information Center of Qinghai Province, Xining 810000, China

*Corresponding author: Xuanyu Zhang, zhangxuanyu0808@163.com

Abstract: Meteorological data is the record and forecast of meteorological changes. It can provide certain guidance for people's production and life, so it is an important information content. Therefore, the current meteorological data management has been transformed from the traditional manual management mode to the business system operating platform management mode, which can better complete the collection, storage, and analysis of meteorological data, and improve the efficiency of meteorological data and data processing. This paper mainly analyzes the role of the operating platform of the meteorological data business system and summarizes the sources of meteorological data misinformation and data quality control strategies to ensure the reliability of meteorological data.

Keywords: Meteorological data business system; Operation platform; Provincial data; Quality control

1. Introduction

With the development of the meteorological industry and meteorological detection technology, the quality control of meteorological observation data has also put forward higher requirements. Since the 1990s, China has undergone large-scale changes in the meteorological data processing business. The current meteorological observation has been fully automated, which can ensure the real-time transmission of data files and provide more accurate meteorological data. The meteorological data business system operating platform was born, realizing the integrated development of meteorological data business management and improving the quality of meteorological data [1]. However, it is still inevitable that there will be wrong information, resulting in the quality of meteorological data being affected. This paper mainly studies the provincial data quality control to form more accurate and precise meteorological data.

2. The function and object of the operating platform of the meteorological data business system

2.1. Functions of the operating platform of the meteorological data business system

The operating platform of the meteorological data business system is an integrated management platform built for the collection and processing of meteorological monitoring data. The platform can form an integrated management mode for meteorological monitoring, collection, processing, and storage of meteorological data, improve the efficiency of data management, and form information sharing among different departments [2]. The construction of the operating platform of the meteorological data business system promotes the information development of meteorological departments and improves the accuracy and reliability of meteorological data. The construction of the operating platform of the meteorological data business system also provides a guarantee for the quality of provincial data, which is directly related to the utilization of meteorological data and the development of related work. If the data quality is not high, it will affect the efficiency of subsequent data use and the accuracy of data publication, reduce the use value of meteorological data, and have a certain impact on the development of other industries [3].

2.2. Provincial data quality control object of the meteorological data business system operation platform

2.2.1. Control of minute data

The control of data quality in the meteorological data business system includes the control of minute data, which is to record the factors such as pressure, temperature, precipitation, wind direction, and humidity in the meteorological environment in minute units, as the main element of meteorological observation [4].

2.2.2. Control of hourly data

The control of hourly data in the meteorological data service system has certain similarities with the control of minute data, including air pressure, temperature, humidity, and so on. The content of hourly data is more abundant, as it also includes radiation, sunshine, ground temperature, grass temperature in deep and bottom layers, and so on. The measurement content is more comprehensive and it is divided into different levels of temperature measurement [5].

2.2.3. Daily data control

In the quality control of meteorological data, the controlling factors of daily data include evaporation value, weather phenomenon, ice accumulation, precipitation, snow depth, and so on. It can be detected by manual observation or instrument detection, and then generate a report on meteorological data [6].

3. Provincial data quality control regulations for the operating platform of meteorological data business system

3.1. province-level data quality misinformation sources

The misinformation in the operating platform of the meteorological data business system is mainly the data information automatically formed by the quality control system in the processing of the platform, and some of the information is the questioning information formed by the relevant personnel in the quality control of the data and the data submission process. In the processing of these types of data, the data that need to be processed

include automatic and manual observation data, metadata, and various statistical data [7]. The quality of minute, hour, and daily data uploaded into the system should be detected and controlled.

3.2. Processing of misinformation about provincial data quality

Before the processing of suspected error data information, it is necessary to quickly retrieve the explicit error information, but the information does not need to be directly modified. First, the information is marked as wrong data information and marked with a quality control code. There are two ways to mark suspected error data in the operating platform of the meteorological data business system, which are error information and suspicious information. Among them, the suspicious information needs to be manually checked at the provincial level and then processed according to the specific situation after confirming that it is correct information or wrong information [8]. If the data information is wrong data information after manual verification, the information should be marked and recorded in the system as metadata. Only quality control should be carried out on the statistical data of changes in atmospheric pressure, temperature, and so on, without data modification. If the manual verification thinks that the statistical data formed in the system is wrong, the relevant data can be corrected to ensure the accuracy of the data [9].

3.3. Provincial data common quality misinformation and processing methods

After the provincial data quality misinformation detection, to better process the misinformation, it can be combined with the classification of meteorological elements to carry out quality control on the misinformation, ensure the quality of the misinformation processing, and prevent the quality control from being affected because of late erroneous operations or inappropriate management processes. Judging from the current provincial data processing situation of the operating platform of the meteorological data business system, most of the suspected error information is reflected in precipitation, cloud cover, weather, ground temperature, and wind speed [10].

3.3.1. The processing of precipitation misinformation

The rainfall hourly data could have misinformation if the suspected description is inconsistent with the spatial consistency check, and the monitoring data of the neighboring station is relatively different. It is necessary to check whether there is local precipitation or problems with precipitation equipment. It is necessary to check the local precipitation and precipitation intensity according to the local weather records and do the inspection work in advance during the weather monitoring process. If there is a large gap between the data that can be described as the total precipitation in minutes and the hourly precipitation, it is necessary to analyze whether there is precipitation lag or non-precipitation data is deleted. It is necessary to determine whether there is a fault in the detection equipment and use the business software to process the minute data to avoid errors in the subsequent minute data processing [11]. In the 24-hour precipitation monitoring, there is a large difference between the precipitation in the hourly data and the actual statistics, so it is necessary to consider the non-precipitation deletion or precipitation lag. In addition, it may be because the data is not uploaded in time. Given this situation, it is also necessary to carry out data processing on the system software platform and correct the influence of the timing data.

3.3.2. Cloud and weather observation processing of suspected error data

If there is a lack of measurement in the detection of total cloud cover and low cloud cover, the hourly data

may have errors. The main reason for this is that the timing data elements and related files are not uploaded to the platform in the system entry. Given this situation, the data can be supplemented and verified by manual input. At the same time, start the standby machine to check the data file, and pay attention to not upload the data repeatedly. In the observation of weather phenomena, if the daily data is missing, for example, there is precipitation in a certain time period, but the corresponding time period is not recorded, the precipitation may not be processed in the time period. In the data processing, the regular precipitation of the daily data can be checked with the actual weather phenomenon, the precipitation situation can be processed, and the data can be transmitted to the platform in time. If the daily data is missing in the observation of weather phenomena, for example, there is a precipitation phenomenon in the time period from 08 to 20, but there is no weather phenomenon about this time period in the system or no precipitation is recorded in the weather phenomenon [12]. It may be because although there was precipitation in this time period, the precipitation had completely stopped at 20:00. Or although there was precipitation in this time period, the subsequent data was not transmitted in time, or the transmission was delayed due to modification. In this case, MOI technology can be used to test the precipitation and weather records of conventional data to determine whether the records between the two are consistent. In addition, if there is an error in the hourly data of a weather phenomenon, for example, the weather phenomenon record contradicts the visibility record [13]. It may be because there was an artificial revision in the weather observation, but there was an error in the visibility revision or the revision was forgotten. In this case, the personnel can enter the correct manual observation data into the system, enter the visibility value of the CW segment in the Z file, and the average value of the 10 minutes (VV segment). Other data should be recorded and processed according to the procedure for missing tests.

3.3.3. The processing of ground temperature suspected error data

Ground temperature refers to the measurement of surface temperature. In the measurement of hourly data, if there is a lack of measurement, the highest temperature and lowest temperature of the surface are displayed blank, it may be because the measured data of the ground temperature sensor is inaccurate due to moisture, or the operation of the ground temperature sensor is faulty, affecting the measurement effect. Given this situation, it is necessary to strengthen the inspection and maintenance of seasonal equipment [14]. The missing data can be replaced by the data of the standby station. For suspected errors in the hourly data in the ground temperature test, the error description is generally positive minute data and hourly data are different. In this case, the data displayed in the system is manually remeasured with the data of the standby station. The error data is corrected and the correct information is uploaded to the system. If the hourly data test of 10 cm ground temperature has suspected error and the continuous test fails, the change amplitude of the data in the 12-hour continuous monitoring is relatively small, or the difference between the historical value and the same period is relatively large. The spatial consistency monitoring of the data is not qualified, and the monitoring results of neighboring monitoring stations are relatively different, which may be due to instrument failure or changes in weather conditions. Given this situation, it is necessary to do a good job of monitoring weather changes, observing and uploading 12 hours of continuous data, and strengthening the judgment and analysis of sunshine, temperature, weather phenomena, and other comprehensive conditions. If there is an hourly data error in the test of the highest and lowest ground temperature, the rapid quality control prompt error information, or the monitoring of the limit value fails to pass may be because the monitoring instrument is faulty, resulting in abnormal data, or the data changes are too large, beyond the scope of conventional quality control. Given this situation, it is necessary to further strengthen the

maintenance of monitoring equipment and pay attention to checking the relevant information of quality control [15].

3.3.4. Processing of wind speed doubtful data

In the processing of wind speed data, if there is suspicious hourly data in the 10 minutes and 2 minutes average wind speed measurement, the suspicious description is that the internal consistency inspection has not passed, or the difference between the 10 minutes and 2 minutes average wind speed inspection is relatively large. Generally, if the difference between the 10-minute and 2-minute average wind speed is too large, with more than 4 meters per second, the equipment may have malfunctioned. The personnel can check the operation of the software system and whether it is prompted to update, and compare the corresponding minute value. If the hourly data of the maximum wind speed is shown as an obvious error, the time described as the extreme value of the wind speed is not within the range of observation. It may be because the measurement data of the wind speed has been manually modified, or the program settings of the station are not reasonable. Given this situation, it is necessary to modify the business software in time and do a good job of collecting and processing the extreme wind speed data. If there is suspicious data in the minute data measurement of wind speed or the wind data does not pass the continuous test, it may be because the weather at that time is light or the wind is quiet.

3.3.5. Processing of pressure error data

In the air pressure monitoring, if the minute data in the air pressure monitoring of the station is suspicious, in which the positive point does not match the hour, it may be because the system data is using the backup station data, so the positive point minute data and the timing data should be verified. If there is a relatively large change value between the data of the previous minute, it may be because of the equipment failure or because the weather changes too fast. It is necessary to pay more attention to the changes in the weather system, closely understand the changes in the system data, and deal with anomalies in time once found. If the sea level pressure data is suspicious in hours, if the internal consistency check is not passed, or if there is a large difference between the air pressure and the temperature value, the temperature value may be replaced, which should be completed by re-setting parameters or calculation. If it is described as missing, the temperature measurement results of the previous 12 hours can be checked, and the integrity of the sea level pressure value can be considered.

4. The provincial data quality control strategy of the operating platform of the meteorological data business system

To ensure the provincial data quality control of the meteorological data business system operation platform, it is necessary to strengthen the management of the data from the source and do a good job in the implementation of various links in the later stage to ensure the accuracy of the data.

4.1. Do a good job in the pre-processing of provincial data quality

It is necessary to pre-process the data before the data acquisition of the operating platform of the meteorological data business system. Before data collection, data transmission and operation as well as data storage should be done to check the data, and data processing, timely find abnormal data and suspicious data, to avoid data omission.

4.2. Strengthen the review of data

Meteorological data business system operating platform in the control of provincial data quality can monitor abnormal data through data review. Given the errors, missing, and duplication of data, timely verify the data and correct the wrong data to ensure the integrity and reliability of the data in the system.

4.3. Pay attention to the quality evaluation of data

In the operation of the meteorological data business system, it is necessary to do a good job in the evaluation of data to ensure that all the data quality can meet the data needs of users. The data quality evaluation can be customized according to the needs of the data. The indicators are evaluated to include the accuracy of data, data integrity, and data collection and processing timeliness.

4.4. Construct the feedback mechanism of data

In the data collection and data processing of the operating platform of the meteorological data business system, once the data quality problems are found, the problems should be recorded and reported in time. Feedback to the relevant management department and verify and correct the suspected error data to avoid the occurrence of suspected error information that may affect the application effect of the later data.

5. Conclusion

In summary, the implementation of provincial data quality control on the operating platform of the meteorological data business system can ensure the accuracy of meteorological data and enhance the value of meteorological data. However, judging from the current data application situation, misinformation still exists widely, and many factors are causing this problem. Personnel can combine the specific error information and the cause of misinformation, take scientific adjustment measures, and do a good job in data management and correction.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Hu W, Xiong X, Wei MD, et al., 2023, Meteorological Data Service Monitoring System and Data Transmission Quality Analysis. Meteorological Hydrology and Oceanographic Instruments, 40(1): 32–33 + 37.

[2] Yang M, Luo G, Tang J, 2023, Application and Quality Control Measures of MDOS in Meteorological Data Business. Automation and Application, 64(16): 7–9.

[3] Lv CY, Wu W, Huang XM, 2023, Analysis and Evaluation of Observation Data Quality of Automatic Surface Meteorological Station in Sichuan Province. Plateau and Mountain Meteorological Research, 43(3): 130–136.

[4] Han B, Zhang M, Su LY, et al., 2024, Design and Implementation of Data Quality Control System for Automatic Weather Station. Science and Technology Innovation and Application, 14(20): 13–16.

[5] Chen XH, Kong LW, Cai YC, 2024, Research on Application Program of Data Quality Evaluation of Regional Automatic Weather Station in Gansu Province. Qinghai Science and Technology, 31(1): 138–143.

[6] Gao ZY, He YF, Yang M, 2023, Processing and Quality Analysis of Aircraft Meteorological Observation Data in

Hangzhou. Meteorological Science and Technology, 51(6): 794–804.

[7] Fan HF, 2022, Development and Application of Real-time Upper-air Meteorological Data Quality Control System. Journal of Ningxia University (Natural Science Edition), 43(3): 298–303.

[8] Zheng YP, Ma Y, Han YQ, 2019, Discussion on Methods to Improve the Availability of Observation Data of Automatic Weather Station. Qinghai Science and Technology, 31(2): 119–122.

[9] Wang SF, Jia YC, Guo XX, 2022, Design and Implementation of Meteorological Data Transmission Monitoring and SMS Alarm System. Yangtze River Information and Communication, 35(7): 79–80 + 83.

[10] Liu X, Zhao XL, Wang L, 2022, Quality Analysis and Enlightenment of Metadata Files of Sichuan Ground Meteorological Stations. China Management Informationization, 2022(5): 184–189.

[11] Peng J, Wang XF, Fu XS, et al., 2023, Observation Quality Evaluation of Millimeter-wave Cloud Radar in Shanghai Area for Operational Application. Chinese Journal of Atmospheric Sciences, 46(3): 466–480.

[12] Long XQ, Cao P, 2019, Study on the Distribution Characteristics of Air Quality Index and its Relationship with Meteorological Factors in 14 Cities in Hunan Province. Agricultural Disaster Research, 13(4): 73–75 + 88.

[13] Gao T, Chen TX, Jiang Y, et al., 2022, Journal of Hainan University (Natural Science Edition), 40(3): 280–286.

[14] Shi CX, Xie ZH, Liu RX, et al., 2023, Research on Multi-source Meteorological Data Fusion Technology and Product Development. Science and Technology Achievements Management and Research, 18(8): 58–61.

[15] Xu Y, 2022, Maintenance and Review Method of Surface Automatic Meteorological Observation Data File. Geological Research and Environmental Protection, 1(4): 119–121.