



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Quality Information Management and Application of AWS

Yan WANG*, Feng GAO, Chao SUN, Yu YU, Yiming LIU

National Meteorological Information Center, Beijing 100081, China

Abstract This paper has designed a unified storage model of Automatic Weather Station Quality Information Inquiries and Feedback (AWSQIF), and has developed an AWSQIF automatic management system, which implements the storage of AWSQIF automatically in national quality information two-way feedback process. Based on the unified storage, AWSQIF data have important application in real-time data management, data quality assessment and quality control effect analysis. The results show that the data quality has been improved significantly, and the data assessment has achieved online operating. In addition, AWSQIF can assist the improvement of the quality control effect.

Key words AWS (Automatic Weather Stations), Quality control, query, Feedback, Quality management

1 Introduction

Since 2009, China has gradually carried out the studies on the real-time quality control methods of basic observation element data of AWS (Automatic Weather Stations)^[1–3], to achieve business transformation. By assessing the hourly observation data and quality control results of national ground automatic stations, it is found that there is a high error rate in automatic station data, leading to decreased availability of information. Taking the second quarter of 2012 for example, the average error rate of 2411 national automatic stations was 0.09%, and the average error rate of 29399 regional automatic stations was 0.77%, that is, the data of 2 national stations and 226 regional stations per hour were not available due to quality problems. In order to grasp the accuracy of quality control results and further improve the quality control algorithms, the query and feedback regarding error data are conducted after the service transformation of quality control system. Since some feedback files are compiled by human, there are many abnormal conditions such as messy code and formatting disorder, making it difficult to effectively assess, manage and apply the feedback data. Meanwhile, the feedback data are not applied in the revised working of real-time observational data and quality information, so forecasting and other real-time service users can not effectively use the data, and can not confirm credibility of some suspicious precipitation, thus affecting the forecasting effect. To improve this situation, the National Meteorological Information Center launched automated management and application of QC (Quality Control) information query and feedback of national automatic stations to further standardize the query and feedback file format of error data, and established national QC information management system; cooperated with Jiangxi, Hubei and Guangxi provincial information center to establish a national and provincial two-way feedback process of national automatic station QC information, and applied

it in 31 provinces to achieve the fully automated query and feedback of quality information. This article focuses on the design and implementation of QC information management system.

2 Overview of national and provincial two-way feedback process

After the quality control over the national AWS observation data by the national real-time data quality control system, the erroneous data, as the original data, are entered into QC information management system, and form query documents to be issued to the province. After the quality control over provincial observational data^[4–5], the data are blended with the national query information received, and by the immediate alarm function of provincial quality control software, the erroneous data information is released to the provincial data processing personnel and weather stations^[6]. After confirmation of the erroneous data, and the province forms feedback files and uploads them to national QC information management system which integrates and stores the feedback information and original query information, and revises the quality control results of real-time automatic station observational data. The overall process is shown in Fig. 1. National and provincial two-way feedback process involves many links such as quality control, quality query and feedback, communication and transmission, which are interrelated. The consistency of data is ensured after passing QC information management system at the national level and passing quality control information fusion subsystem at the provincial level.

3 System design and information flow

3.1 System structure design The overall system is designed as three-layered structure (data interface, data processing and data storage), as shown in Fig. 2. With the national storage and retrieval system MDSS real-time database as data environment, the data are stored and the quality management table AWSR is designed to store the query and feedback information of erroneous data. As the data interface for various links in national and provincial two-way feedback process, it determines the format convention

Received: November 18, 2015 Accepted: December 28, 2015

Supported by Business Construction Project of China Meteorological Administration in 2011 and Key Project of Key Meteorological Technology Integration and Application of China Meteorological Administration (CMAGJ2013Z01).

* Corresponding author. E-mail: wangyan8@cma.gov.cn

of erroneous data, query files and feedback files. Data processing consists of three modules, namely national QC erroneous data storage, national query generation and feedback data storage. (i) National QC erroneous data storage. The erroneous data from the national quality control system are stored as the data source of QC information management system. The storage recording data include station number, observation elements, observation time, observational value of elements, quality control results and other query information, in conjunction with data state, storage time and other process information. (ii) National query generation. The file query is conducted based on unreleased records in AWSR table. The query files are generated in the light of province, and the files are placed in the distribution directory of communication system which transfers the query files to the provincial level. In the process of generating query files, the data state is modified in order to avoid interference from other processes. (iii) Feedback data storage. The feedback data are generated from the content of feedback files one by one and input into database in real time. The query information contained in feedback data is used for positioning the corresponding query record that already exists in AWSR table, and updating the feedback information, feedback time and data state. For the query record that does not exist in the table, it is directly inserted. QC information management is to conduct unified storage

and comprehensive application of query and feedback data from multiple sources. Based on such unified storage, we can revise and update the real-time observation data and quality information, and analyze the real-time feedback and quality control results. The confirmation results of feedback data is of great significance to quality information of automatic station observation data. After completing the storage work of AWSR table, the confirmation results of feedback data are directly applied to automatic station hourly data sheet, to revise the quality control code of corresponding observational data. In terms of scheduling configuration, the state-level quality control system takes the multiple quality control processing mode based on scheduling strategy^[7], and the quality control results will change over time, so the QC information management system needs to reserve the final erroneous data and store them into database after processing all the results performed within one hour by quality control system. There is successive relationship between erroneous data storage and query generation, and erroneous data storage can be only performed after the end of the current hour, so the two functional modules are combined and started at the fifth minute of next hour after the current hour quality control is completed through the regular scheduling. The feedback data storage is not limited, and it is regarded as the permanent system process to carry out real-time processing of feedback files.

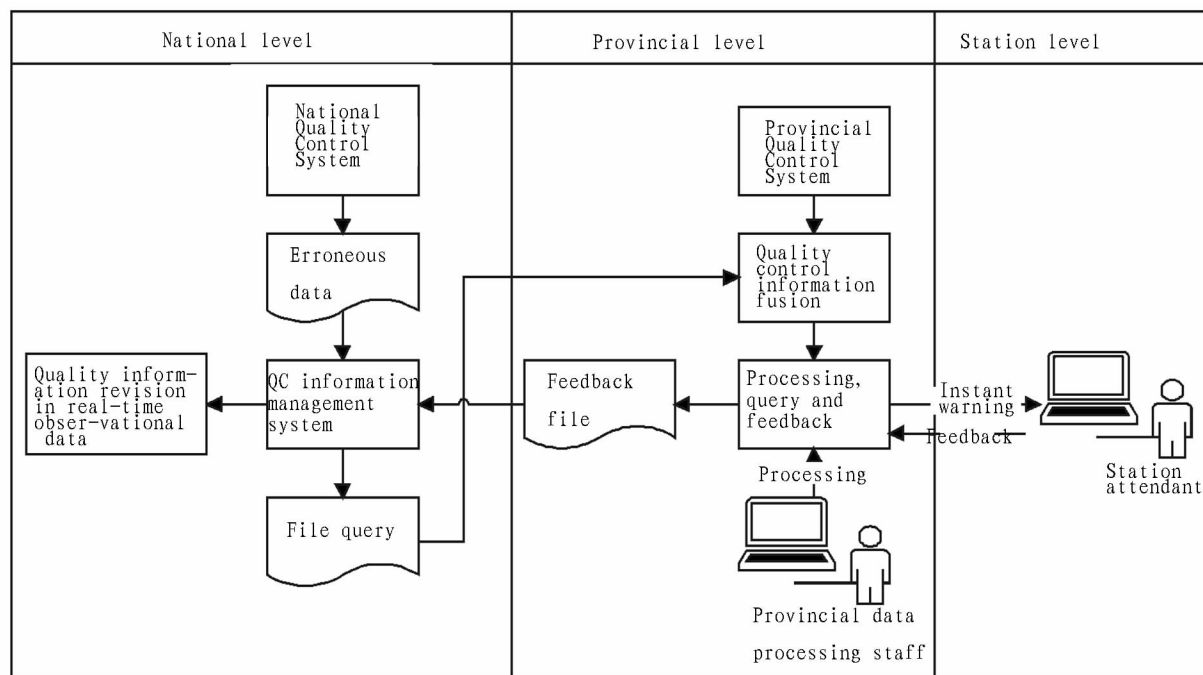


Fig. 1 National and provincial two-way feedback process

3.2 "Asynchronous" information exchange mechanism

The national automatic station QC management system consists of three separate sub-processes: national QC erroneous data storage; national query generation; feedback data storage. With AWSR table as a core, the "asynchronous" information exchange mechanism is used and the data state is marked to achieve convergence

and interaction of three sub-processes. After the completion of storage, query and feedback processes, the corresponding data state is identified as "having not issued", "having issued" and "having fed back", and "having fed back" is the final data state. In the real situation, storage always occurs prior to query, while feedback generally follows query, but may also occur before and

after storage. (i) Storage-query-feedback. It is one of the most common business processes. After the national quality control, the erroneous data are stored and the query data files are formed to be issued to the province by the communication system. The data state is "having issued". After confirming data quality, the provincial data processing personnel perform the feedback, and data state changes to "having fed back". It is often based on the last feedback if the feedback occurs many times. (ii) Storage-feedback. The erroneous data from provincial quality control system would also produce query information, and the QC information management system receives the provincial query and feedback data at the same time. When the erroneous data are stored and have not been issued, the provincial feedback information has been put

in storage, the data state changes to "having fed back", and the query sub-process will no longer be executed in accordance with such data state. (iii) Feedback-storage. The national and provincial quality control results are not exactly the same, and when the provincial query and feedback data are stored, there may be no corresponding national query data to be stored. The data state is "having fed back". In order to distinguish query information sources, the auxiliary mark of "query sponsor" is added. If the national erroneous data are stored after the provincial query and feedback data, it will not change the data state of "having fed back", feedback results and query sponsor information, but only update the storage time of national erroneous data.

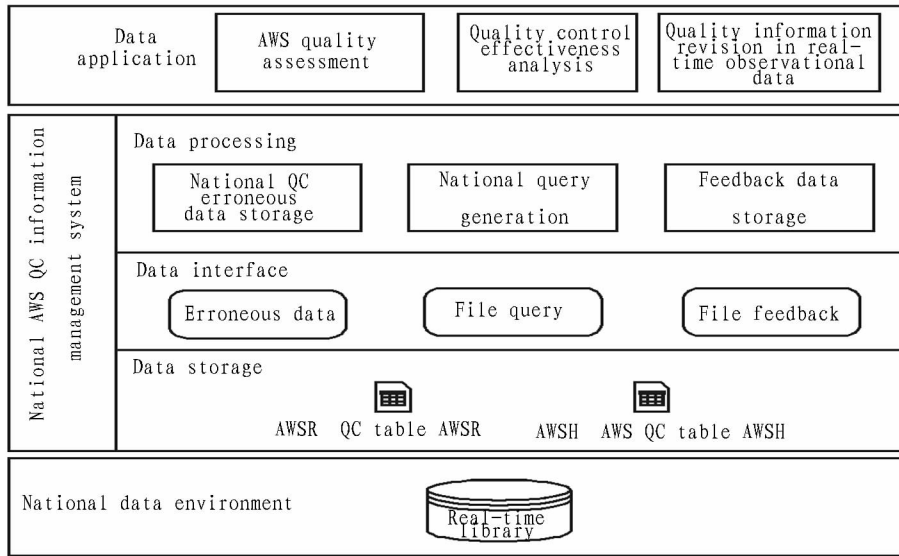


Fig.2 Structure of QC information management system

3.3 Database design The erroneous data contain the basic query information, namely station number, observation elements, observation time, observation value and quality information. For feedback information, it includes the confirmation of data quality, cause of problems and note information. Seeing query and feedback as entity, a piece of query data is allowed to be fed back many times, but it is based on the last piece of feedback data, that is, a piece of query data corresponds to a piece of feedback data at most. For a piece of feedback data, there must be a piece of query data. Therefore, there is one-to-one relationship between query entity set and feedback entity set. According to the simplicity principle of relational data model design, the two types of entity set are designed as a unified relational data model AWSR, as shown in Table 1. In order to ensure the correct interaction of storage, query and feedback sub-processes and effective management of storage data, the flow information fields are added, including data state, query sponsor, query storage time, and feedback time.

4 System construction and implementation

4.1 Deployment operation QC information management sys-

tem was deployed and put into operation nationwide in July 2012. In the system, the feedback data can revise the real-time data quality and modify quality control identification of observational data based on confirmation results in real time as: correct (0), erroneous (2) or can not be confirmed (1). According to the hourly observational data and quality control assessment results of national automatic stations, it is found that the average error rate of national stations was 0.002% and the average error rate of regional stations was 0.07% from July to November 2012. If the quality feedback data is not applied to revision of real-time data quality, the average error rate of national stations is 0.009% within this period, and the average error rate of regional stations is 0.10%. Fig. 3 shows the change in error rate before and after quality information revision of national stations and regional stations from July to November, and we can find that the error rate of automatic stations after being revised each month is significantly lower than before being revised, which means that the revision and application of QC information management system has improved the availability of data. By establishing the query and feedback two-way process, the national or provincial stations can grasp the

data quality in a timely manner, and timely solve the data problem caused by various anomalies (such as equipment failure); the data error rate of automatic stations during July to November fell

sharply compared with April to June, and the data quality has improved significantly.

Table 1 Storage structure design of AWSR

No.	Field name	Data types	Field description	Classification
1.	queryID	CHAR(30)	Observation time, station number, element code combination	Primary key
2.	qinsert_time	CHAR(14)	National query storage time, default value ' - '	Flow information
3.	fcreate_time	CHAR(14)	Feedback file generation time, default value ' - '	
4.	frev_time	CHAR(14)	Receiving time of feedback file, default value ' - '	
5.	finsert_time	CHAR(14)	Feedback storage time, default value ' - '	
6.	sponsor	VARCHAR2(30)	National quality control sys; ID of users questioned; provincial query prvc	
7.	state	NUMBER(1)	0;having not issued; 1;having issued, yet to be fed back; 2;having issued; 9;in the process of being issued	Query information
8.	station	NUMBER(10,4)	Cannot be null	
9.	code	VARCHAR2(20)	Cannot be null	
10.	year	NUMBER(4)	Cannot be null	
11.	month	NUMBER(2)	Cannot be null	
12.	day	NUMBER(2)	Cannot be null	
13.	hour	NUMBER(2)	Cannot be null	
14.	minute	NUMBER(2)	Cannot be null	
15.	second	NUMBER(2)	Cannot be null	
16.	value	NUMBER(2)	Cannot be null	
17.	flag_q	CHAR(1)	Default value ' - '; 1 = suspicious, 2 = erroneous	Feedback information
18.	flag_f	CHAR(1)	Default value ' - '; 0 = correct, 2 = erroneous, 1 = cannot be confirmed	
19.	reason	CHAR(1)	Problem reason code, value ranging from A to F, default ' - '	
20.	remark	VARCHAR2(100)	Chinese character string, default null	

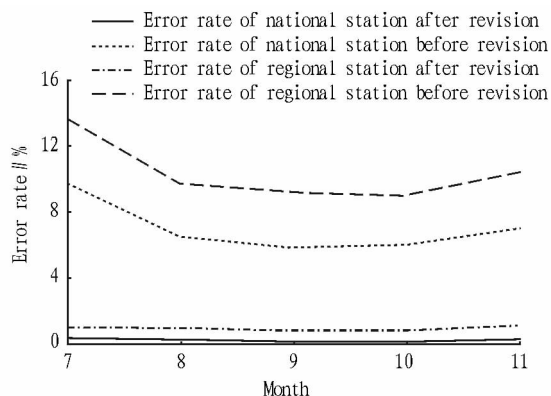


Fig. 3 The average error rate comparison of national automatic stations before and after quality revision during July to November 2012

4.2 Statistical analysis of feedback rate The national automatic station quality assessment business is one of the main applications of AWS quality information, and the specific assessment items involve feedback rate, amount of correct and incorrect suspect data feedback, amount of correct and incorrect erroneous data feedback, etc. The feedback rate is the ratio of feedback amount of data to query amount of data, and feedback amount of erroneous data is counted by the element quality mark (flag_q) and confirmation results (flag_f). AWSR table records the query feedback

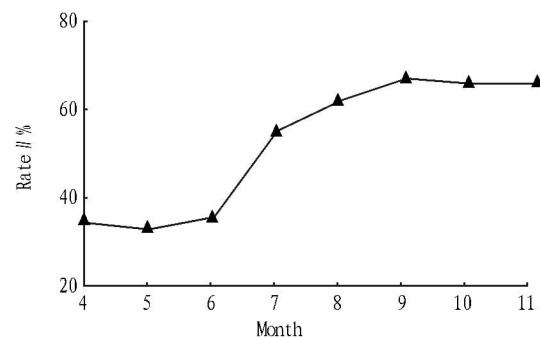


Fig. 4 The average feedback rate of erroneous data of national AWS from April to November in 2012

data content and process information in detail, providing the statistical basis for the feedback rate calculation. All counting indicators can be easily converted to SQL statements to directly complete counting. Given the convenience and operability of this statistical method, we can make automated transformation about it to achieve statistical functions in real time or at any time. Fig. 4 shows the average erroneous data feedback rate of national automatic stations from April to November 2012, the feedback rate from April to June was the statistical value obtained by analyzing and counting the old feedback files, and the feedback rate after July 2012 was the statistical value by statistical analysis of AWSR

table. Due to different making methods of old feedback files in various provinces, there are some non-standard file formats and messy codes, leading to inadequate statistical analysis. Since the system business was started in July, the effective recording rate of feedback files has reached 100%, and from the statistical results of feedback rate, it can be found that the average feedback rate during July to November is 29.08% higher than during April to June, with significant effect.

4.3 The application effect of feedback information The quality information is of important application value to the analysis of quality control effect. First of all, only when the erroneous or suspicious data are proved to be correct or wrong can it be effective feedback, otherwise it will be ineffective feedback. According to the statistics, during the flood season from July to September in 2012, it was judged that the amount of erroneous or suspicious basic elements (such as temperature and hourly precipitation) was 167257 by quality control, with effective feedback amount of 110424 and effective feedback rate of 66%. The effective feedback amount and effective feedback rate of various elements are shown in Table 2, and it can be found that the effective feedback rate of various elements is more than 50%, and we can use the effective feedback information to further analyze the quality control effect. Table 3 shows the QC effect indicators of the basic elements from July to September in 2012, including misjudgment rate of erroneous data, accuracy rate of misjudged data, accuracy rate of suspicious data and error rate of suspicious data. For any kind of element, the sum of the above four indicators is 1. The effective feedback rate of erroneous data is the sum of misjudgment rate of erroneous data and accuracy rate of misjudged data; the effective feedback rate of suspicious data is the sum of accuracy rate of suspicious data and error rate of suspicious data. According to statistics, the misjudgment rate of erroneous temperature and relative

humidity data is low, and the accuracy rate is high, indicating that the quality control algorithm of the two elements plays a significant role in judging the erroneous data. The difference between the accuracy rate and error rate of suspicious data is only about 10%, without the problem of too high accuracy rate or error rate, indicating that the quality control algorithm also plays a significant role in judging the suspicious data. For hourly precipitation, 2-minute wind direction and 2-minute average wind speed, the effective feedback rate of erroneous data is significantly lower than that of suspicious data, while it is opposite for the air pressure element of the station. For the first three elements, the feedback of suspicious data has greater reference value, and it is found that the accuracy rate is 20% higher than the error rate. To further clarify the reasons for high accuracy rate of suspicious data, we can use the reason and remark in feedback data for auxiliary analysis. Taking precipitation for example, in 33505 pieces of precipitation query data, 23746 are fed back, and 2122 pieces of feedback data contain remark information. Based on the analysis of remark information, it is found that 60% of remark content suggests that due to heavy precipitation, thunderstorms and strong convection weather in local areas, the actual precipitation is correct, and the data containing such remarks indicate that the suspicious data feedback is correct. For example, in Shaanxi Huayin (57055), the precipitation was 55.9mm at 9:00 on July 4, 2012; it is judged to be suspicious by quality control system; the remark information is "storm". If we effectively judge local precipitation based on weather radar and other sources of information, the suspicious precipitation rate will be significantly reduced, thereby improving the usability of local heavy precipitation data. Therefore, the information of feedback data helps to find typical weather or problems and constantly improve quality control effect.

Table 2 QC feedback of basic elements from July to September in 2012

Element name	Erroneous data amount	Effective feedback amount	Effective feedback rate//%
Hourly precipitation	35778	24654	68.91
Air temperature	59074	37244	63.05
Air pressure of the station	48369	32493	67.18
2-minute wind direction	2507	1959	78.14
2-minute average wind speed	12668	7443	58.75
Relative humidity	8861	6631	74.83

Table 3 QC effect indicators of the basic elements from July to September in 2012

Element name	Misjudgment rate of erroneous data//%	Accuracy rate of misjudged data//%	Accuracy rate of suspicious data//%	Error rate of suspicious data//%
Hourly precipitation	1.77	1.92	76.06	20.25
Air temperature	4.27	35.91	39.07	20.75
Air pressure of the station	64.67	32.30	2.23	0.80
2-minute wind direction	10.06	18.12	46.25	25.57
2-minute average wind speed	0.50	8.68	73.96	16.86
Relative humidity	10.12	49.95	25.85	14.08

5 Conclusions

National AWS QC information management system is a state-level core system having a national and provincial two-way feedback process.

With national real-time database as storage environment, it designs unified storage table structure, and achieves "asynchronous" information (To page 57)

education of community residents about environmental protection and forest experience, and guide community residents to develop forest experience activities with local characteristics, to promote common development of Forest Park and communities.

6 Conclusions

We put forward development strategies for forest experience from concept, theme, products, ecology, and community, and came up with basic requirements and methods from three levels. The first level takes forest experience concept as guidance and sticks to forest experience management; the second level takes theme and products as the core and provides methods at technical level of forest experience products; the third level is to establish harmonious relationship between forest experience development, ecology and communities, to realize sustainable ecological, social and economic development. These development strategies not only enrich theoretical foundation of forest experience of West Mountain Forest Park, but also satisfy forest experience demands and development experience development methods of Forest Park, and realize the integration of development of Forest Park, social demands, and ecological benefits. It is expected to provide certain guiding significance for multi-functional use and sustainable development of West Mountain Forest Park, and provide some references for forest experience development of other forest parks.

References

- [1] PINE BJ, GILMORE JH. The experience economy[M]. Boston, Massachusetts, Harvard Business School Press, 1999:1–15.
- [2] PI PF, ZHANG WQ. A study on tourism consumption and market strategy based on experience economy[J]. Journal of Jiangxi University of Finance and Economics, 2005(3): 64–67. (in Chinese).
- [3] Bandyopadhyay R. “To be an Englishman for a day”: Marketing colonial nostalgia in India[J]. Annals of Tourism Research, 2012, 39(2): 1245–1248.

(From page 52)

tion exchange of storage, query and feedback based on the data state information in table to ensure data consistency. The real-time revision of feedback information on real-time data quality has significantly improved the availability of observational data. After the establishment of national and provincial two-way feedback process, the feedback rate growth of erroneous data query is very obvious. At the same time, through statistical and visual analysis of quality control effect indicators, combined with feedback data content, the quality control effect can be improved. Since the data interface design is standardized and expandable, QC information management system has been applied in real-time integrated business, to carry out the total factor quality information management of ground data. The next step will be to consider the integrated quality information management of various types of meteorological data.

- [4] WANG F. Research on development of experiencing tourism of Qiaos courtyard based on ASEP grid analysis[J]. Journal of Northwest Sci-Tech University of Agriculture and Forestry(Social Science), 2011, 11(1): 85–88. (in Chinese).
- [5] YANG AL, NAN Y. The development and innovation of the tourism products in the Silk Road based on the viewpoint of the experience economy[J]. Journal of Arid Land Resources and Environment, 2011, 25(8): 196–200. (in Chinese).
- [6] GROSS MJ, BROWN G. An empirical structural model of tourists and places: Progressing involvement and place attachment into tourism[J]. Tourism Management, 2008, 29(6): 1141–1151.
- [7] WEI XH. Management and application pattern for multifunctional forest[D]. Beijing: Beijing Forestry University, 2013: 5–17. (in Chinese).
- [8] WANG JF. A review on forest multi-function management[J]. Forest Inventory and Planning, 2013, 38(4): 131–136. (in Chinese).
- [9] CHENG XP, WANG YF, GONG HD, *et al.* Study on the model of forest multifunctional allocation based on social requirements[J]. Journal of West China Forestry Science, 2014, 43(4): 112–116. (in Chinese).
- [10] WU ZW, WU CC, WEN SW. Forest tourism[M]. Beijing: China Travel & Tourism Press, 2008: 1–3. (in Chinese).
- [11] ZHENG YX, ZHOU FG. Experience-oriented scenic area development model: A new tourism development ideal[J]. Human Geography, 2007, 22(6): 16–20. (in Chinese).
- [12] YU DX, LIU XS. Study on the tourism development of suburban forests from the perspective of experience economy[J]. Forest Resources Management, 2011(2): 21–26. (in Chinese).
- [13] HILLR, PINE BJ, GILMORE JH, *et al.* Welcome to the experience economy[J]. Harvard Business Review, 1998, 44(5): 97–105.
- [14] CHENG XP, CHEN XF, YE W, *et al.* Growth of forest experience in Japan and its revelation[J]. World Forestry Research, 2015, 28(2): 75–80. (in Chinese).
- [15] TRAUER B, RYAN C. Destination image, romance and place experience: An application of intimacy theory in tourism[J]. Tourism Management, 2005, 26(4): 481–491.
- [16] ZOU TQ, HAO LY. The nature and categories of tourism experiences and developing principles[J]. Tourism Science, 2003(4): 7–41. (in Chinese).
- [17] LIU JM, LIU Y. An analysis of tourism revitalization of historic district in perspective of tourist experience: A case study of San-Fang Qi-Xiang district in Fuzhou City[J]. Geographical Research, 2010(3): 556–564. (in Chinese).
- [18] HU YW, ZHANG P. Study on tourism development in the age of experience economy[J]. Journal of Guilin Institute of Tourism, 2003, 14(3): 48–51. (in Chinese).

References

- [1] REN ZH, ZHAO P, ZHANG Q, *et al.* Quality control procedures for hourly precipitation data from automatic weather stations in China[J]. Meteorological Monthly, 2010, 36(7): 123–132. (in Chinese).
- [2] ZHAO YF, REN ZH, ZHANG Q. Quality control procedures for hourly relative humidity data from national automatic weather stations[J]. Scientia Meteorologica Sinica, 2011, 31(6): 687–693. (in Chinese).
- [3] JU XH, REN ZH, ZHANG Q. Research on quality control technology of automatic hourly pressure[J]. Journal of Anhui Agricultural Sciences, 2010, 38(27): 15130–15133. (in Chinese).
- [4] DOU YW, QU YG, TAO SW, *et al.* The application of quality control procedures for real-time data from automatic weather stations[J]. Meteorological Monthly, 2008, 34(8): 77–81. (in Chinese).
- [5] WANG HJ, YANG ZB, YANG DC, *et al.* The method and application of automatic quality control for real time data from automatic weather stations[J]. Meteorological Monthly, 2007, 33(10): 102–109. (in Chinese).
- [6] LI ZP, ZHANG W, HUANG SP, *et al.* Design and implementation of the software for real-time quality control over automatic weather station observation data[J]. Meteorological Monthly, 2012, 38(3): 371–376. (in Chinese).
- [7] LIU YM, ZHOU ZJ, YUAN F, *et al.* ARQBS starting strategy and its relationship with computing resource cost[J]. Quarterly Journal of Applied Meteorology, 2014, 25(4): 505–512. (in Chinese).