

李志鹏,张玮,黄少平,等. 自动气象站数据实时质量控制业务软件设计与实现[J]. 气象,2012,38(3):371-376.

# 自动气象站数据实时质量控制 业务软件设计与实现<sup>\*1</sup>

李志鹏 张 玮 黄少平 李洪康 邓卫华

江西省气象信息中心,南昌 330046

**提 要:** 介绍了一种面向天气监测预报实时业务的自动气象站数据质量控制软件。该软件综合应用数据质量控制、网络消息即时通信等技术,将数据质量控制纳入气象信息业务流程,建立了省、市、县三级气象台站联动的自动站数据实时质量控制业务流程,实现了对江西省 89 个国家自动气象站和 1531 个区域自动气象站观测数据的实时质量控制,有效降低了自动站疑误数据对天气监测预报业务服务可能造成的不利影响。

**关键词:** 自动气象站, 疑误数据, 实时预审, 质量控制

## Design and Implementation of the Software for Real-Time Quality Control over Automatic Weather Station Observation Data

LI Zhipeng ZHANG Wei HUANG Shaoping LI Hongkang DENG Weihua

Jiangxi Meteorological Information Center, Nanchang 330046

**Abstract:** A real-time weather monitoring and prediction operation oriented software for quality control over automatic weather station (AWS) observation data is introduced. The software makes comprehensive use of data quality control method and instantmessaging technology, brings data quality control into meteorological information operation flow and thus a province-city-county three level real-time quality control operation flow over AWS observation data is built. The observation data from 89 national AWS and 1531 regional AWS in Jiangxi Province are under real-time quality control. The adverse effect on weather monitoring and prediction caused by AWS questionable data is effectively reduced.

**Key words:** automatic weather station (AWS), questionable data, real-time pre-examination, quality control

## 引 言

由于受周边探测环境及自动站自身软硬件稳定性等多种因素影响<sup>[1]</sup>,自动站采集的观测数据可能会出现非真值,这些失真的数据进入预报业务系统,将可能误导对天气系统的监测分析和预报服务。如何实时高效地开展自动站数据质量控制,保证输入到预报业务系统资料的准确有效,是气象信息系统

研究的一项重要课题。

针对自动站数据质量控制,任之花等<sup>[2-3]</sup>对非实时资料做过质量控制,王海军等<sup>[4]</sup>、窦以文等<sup>[5]</sup>对省级收集的自动站资料做过实时质量控制,王建庄等<sup>[6]</sup>对采集器数据做过实时质量控制。但是,把包括区域自动站(以下简称“区域站”)在内的数据质量控制纳入实时业务流程未见文献报道。

自动站观测数据作为天气预报特别是灾害性天气监测预警的重要基础资料,对其的基本要求是及

\* 江西省气象科技创新基金项目“省、市、县三级自动站疑误数据快速处理反馈业务系统研究”资助  
2011年4月23日收稿; 2011年9月29日收修定稿  
第一作者:李志鹏,主要从事气象信息系统业务与应用研究. Email: lzp2009@tom.com

时、有效。及时:主要体现在资料的实时性上,要求做到数据即采即用,观测数据能在最短时间内进入预报业务系统;有效:主要体现在资料的准确性上,要求观测数据尽可能真实反映大气的实际物理状态。本文所介绍的自动站数据实时质量控制软件,针对天气监测预报实时业务,集数据实时预审、疑误数据监视告警、数据修订反馈等于一体,实现了对省级收集的自动站数据即到即审,并基于网络对出现可疑数据的测站(气象台站)即时定向告警提示,动态生成测站可疑数据检查分析所需基本资料,基于 Web 在线反馈检查结果、查询疑误数据的出现与处理情况,从而实现省、市、县三级自动站数据的实时质量控制。

## 1 系统结构与数据流程

### 1.1 系统结构和功能

系统结构与数据流程如图 1 所示。考虑到系统的稳定性与执行效率,软件采用了 C/S 与 B/S 混用的架构方式,运行于省以下气象广域网络环境,其中,服务器端部署于省级中心,实现数据实时预审、即时告警信息的生成与发送功能(C/S),并负责向台站提供分析检查可疑数据所需的基本资料、接收台站数据检查修订结果(B/S);客户端部署于各气象台站,实现可疑数据告警信息的监测与显示功能

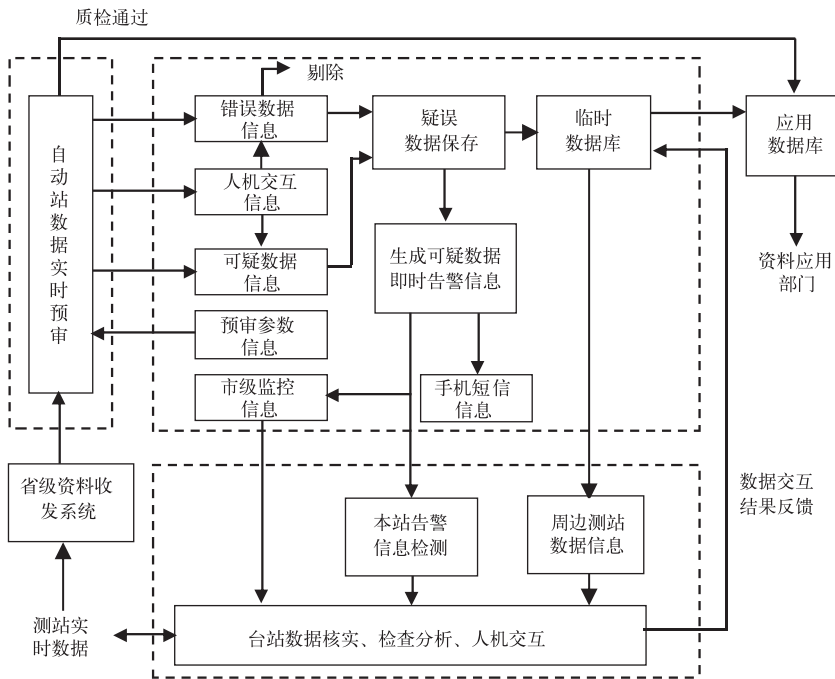


图 1 省、市、县三级自动站数据质量实时控制系统结构与数据流程

Fig. 1 Systematic organization and data flow of province-city-county three level real-time quality control over AWS observation data

(C/S),并负责本地可疑数据的分析检查与修订反馈(B/S)。在台站客户端,B/S 应用部分设计为一个功能模块,嵌入在 C/S 模式的主体程序中。

从应用功能划分,系统主要包括观测数据实时预审、可疑数据即时告警、可疑数据处理三个部分。

#### 1.1.1 观测数据实时预审

主要采用气候界限值检查、极值检查、内部一致性检查等方法<sup>[7-8]</sup>,对实时收集的自动气象站观测数据(包括温度、降水、气压、风、湿度、地温等要素)进行随到随审,甄别疑误数据并保留其原始数据文件,

形成临时数据库。

#### 1.1.2 可疑数据即时告警

系统一旦监测到某站出现可疑数据,将自动触发对该站的定向告警,同时,部署于该站的客户端计算机桌面将立即弹出多媒体告警信息窗口(如图 2 所示),可疑数据也随告警信息一并返回到气象台站。如告警信息未得到及时处理,告警将一直持续;如告警信息被关闭,系统将每隔一定时间(由服务器端设置)启动告警;若台站连续多个时次出现告警而未处理,这些告警信息将始终有效,直至所有告警信

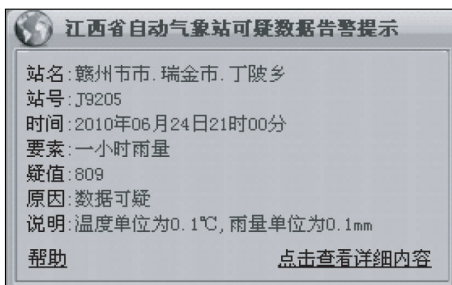


图 2 台站可疑数据告警提示窗口  
Fig. 2 The pop-up window alerting questionable data



图 3 可疑数据所在测站周边站点数据显示  
Fig. 3 The display of observation data surrounding questionable data station

核查修订结果。

### 1.2 数据流程

如图 1 中,系统部署于省级实时资料收发系统与自动气象站应用数据库之间。省级实时收集的国家自动站(以下简称“国家站”)和区域站正点观测资料作为该系统的数据输入,系统的输出是自动站应用数据库的数据源。

依据预先设置的数据预审规则,由自动站数据实时预审软件自动对数据进行质量检查(软件自动预审中难以界定真伪的要素值则辅以人机交互检查)<sup>[9]</sup>,剔除错误数据,标记可疑数据,质检结果为正确的数据则直接进入自动站应用数据库。

根据可疑数据标记,系统生成可疑数据即时告警信息对远端台站定向告警,台站客户端实时检测省级中心服务器对本地的可疑数据告警信息,业务人员通过调用系统提供的数据分析检查所需基本资料,对可疑数据核查分析修订。核查修订后的数据采用网页表单递交方式(B/S)向省级中心服务器反馈。

由于区域站无人值守,对应的可疑数据告警信息发送到该区域站所属台站,并由其负责处理。必

息处理完毕,每条告警信息独立处理,互不影响。

#### 1.1.3 可疑数据处理

在向台站返回可疑要素测值的同时,省级中心服务器端还从临时数据库中获取该测站周边±0.2经纬度范围内其他测站的同时次数据,并与相关的公路、铁路、河流等地理信息集成,动态生成以该可疑数据所在测站为中心的临近站点要素标注填图提供给台站,以便于可疑数据的分析和判断(如图 3 所示)。台站根据本地天气实况和邻近站参考数据,经过必要的对比、检查、核实和分析后,反馈可疑数据

要时台站还需与区域站所在地兼职维护管理人员(或气象信息员)联系,进行进一步的检查核实。

为提高可疑数据处理的及时性,并加强市级业务管理人员的督促与技术指导作用,系统同时以手机短信方式向市级业务人员发送所辖各台站测站可疑数据告警提示信息,及向县级发送所辖测站可疑数据告警提示信息,系统还于每日 08 和 17 时向市级业务管理人员短信通报前一时间段所辖各台站可疑数据处理情况。通过全省气象内网,省、市两级及各县级台站可对疑误数据的发生及处理情况进行实时监测与查询。

## 2 关键技术

### 2.1 区域站数据的实时预审技术

区域站数据的实时预审主要采取气候界限值检查、历史极值检查和空间一致性检查三个步骤实现。

#### 2.1.1 极值和界限值检查

由于区域站建站时间不长,无法形成有效的极值序列,因此区域站极值以邻近的国家站历史极值

为基础,根据其所处位置、探测环境等进行订正后作为该区域站历史极值,即:区域站历史极值=邻近国家站极值±2 倍的国家站极值序列标准差±探测环境订正值。

区域站界限值的确定则是在国家站气候界限值的基础上,采用极值序列标准差的 10 倍作为界限值的参考订正依据,结合有可能出现的特殊天气(如台风、龙卷风等)对气象要素的影响,生成江西常用要素气候界限值指标。气压、温度按月设置极值和界限值范围,风速、降水等按年设置极值和界限值范围。以下为常用要素的极值及界限值的计算方法。

(1) 气压:将区域站本站气压订正到海平面气压(地形特殊或高海拔地区可以考虑将气压订正到某一高度),以此进行极值和界限值检测,同时也便于空间一致性检测。具体方法是通过对国家站逐年同期出现的极值订正到海平面值,计算最高(低)值序列的标准差后得到气压历史极值,考虑到可能出现龙卷风等天气,界限值下限维持为 400 hPa。海平面气压计算公式为:

$$p_0 = p_h \times 10^{\frac{h}{18400(1+\frac{t_m}{273})}}$$

式中, $p_0$  为海平面气压(单位:hPa); $p_h$  为本站气压(单位:hPa); $h$  为气压传感器(水银槽)拔海高度(单位:m); $t_m$  为气柱平均温度(单位:°C),计算气柱平均温度  $t_m$  公式:

$$t_m = \frac{t+t_{12}}{2} + \frac{\gamma h}{2} = \frac{t+t_{12}}{2} + \frac{h}{400}$$

式中, $t$  为观测时的气温(单位:°C); $t_{12}$  为观测前 12 小时的气温(单位:°C); $\gamma$  为气温垂直梯度或称为气温直减率,规定采用  $0.5^\circ\text{C} \cdot (100\text{ m})^{-1}$ ;  $h$  为气压传感器(水银槽)拔海高度(单位:m),对于一个测站来说, $h$  是一个定值,故  $h/400$  为一常数。

(2) 气温:由于安装在平顶屋面自动站较安装在标准场地的区域站在晴好天气条件下最高温度可高  $3\sim 4^\circ\text{C}$ ,最低温度低  $1\sim 2^\circ\text{C}$ ,其差异值约为标准场地极值的 10%左右,故温度探测环境订正值=国家站气温极值×10%。

(3) 小时降水:从江西省 1980 年以来年报表和降水日记数字化成果数据中挑取了出现过的 1 小时最大降水量:132.6 mm(庐山站,1998 年 9 月 15 日),序列的标准差为 23.9 mm,将区域站小时降水的极大值设为 180 mm,小时降水界限极大值设为 370 mm。

(4) 风速:现有记录逐年序列中出现的最大风极值: $30.0\text{ m} \cdot \text{s}^{-1}$ (泰和站,1977 年 4 月 24 日 17 时 10 分),序列的标准差为 3.0,将区域站 10 分钟

平均风速的极大值设为  $36\text{ m} \cdot \text{s}^{-1}$ ,界限极大值设为  $60\text{ m} \cdot \text{s}^{-1}$ 。

### 2.1.2 空间一致性检测

对现有各区域站(间距 5~50 km 范围)逐时观测资料统计分析表明:随着距离的增加,两站之间气压和温度的相关系数变化不大;而降水和风速的相关系数则明显下降。上述四种要素在站址相距 20 km 范围内时均具有较高的相关性(见表 1),除此之外,再考虑到中小尺度天气系统的空间尺度等因素,系统选择了台站周边±0.2 经纬度(约 22 km)范围内各测站数据作为进行空间一致性检查的参考序列<sup>[10-11]</sup>。

表 1 区域站站间气象要素相关系数  
Table 1 Statistics of correlation coefficients of meteorological elements in regional AWS

距离	气压	气温	相对湿度	降水	风速
20 km 内	1.000	0.996	0.951	0.670	0.636
30~37 km	0.999	0.992	0.924	0.541	0.568
40~50 km	0.999	0.993	0.910	0.403	0.532

台站业务人员通过人机交互将可疑数据与其周围其他邻近测站的测值(包括本时次及过去若干时次数据)进行比较分析,结合本站实况、雷达回波、自动站设备运行等,判断该要素是否正常<sup>[12]</sup>。这种方法虽属定性判断,但对降水数据的质量控制具有实用价值,尤其是出现较强降水数据而又难以进行定量判断时,该方法有明显作用。

## 2.2 可疑数据告警信息的侦测与处理

告警信息由服务器端主动发送,台站客户端依靠常驻内存的守护进程,完成对告警信息的侦测、显示和数据修订。

### 2.2.1 告警信息侦测

由于可疑数据的出现具有随机性,台站客户端需要不间断地检查指向本站的告警信息,为了避免受到客户端其他应用所产生的终端信息干扰,告警信息侦测宜通过守护进程实现,该守护进程在计算机启动时自动运行,可长时间运行于特定用户帐户下,通过 OnStart 和 OnStop 方法为其指定代码以定义服务的行为方式。在客户端确认收到告警信息后,系统以 GetCurrentProcess 方法回收系统内存,保证客户端系统内存占用最小化<sup>[13-14]</sup>。

### 2.2.2 数据告警提示

客户端守护进程通过 sock 方式调用可疑数据告警接口,数据告警接口直接调用 Web Service 中间层组件实现告警信息传递。当告警事件发生时,由 SoapClient 生成 Soap 请求消息发往 SoapServer,

SoapServer 接收到 Soap 请求消息后,完成指定功能并把返回值(该站出现的具体疑误数据)放入 Soap 回应消息,返回给 SoapClient,由台站客户端对该疑误数据进行提示告警<sup>[15]</sup>。

### 2.3 基于 flash RIA+asp.net+xml 的邻近测站观测数据显示

邻近测站观测数据显示处理流程如图4所示。传统的 web 应用是基于 HTML 页面的服务器端数据传递的应用模式,而互联网应用 RIA(Rich Internet Applications)则将桌面应用与传统的 Web 应用融合为一体,其既具有 Web 应用部署的灵活性,又具有桌面应用强大的计算能力和交互能力。RIA 使得客户端可以显示和操作更为复杂的数据模型,从而提高与用户的交互能力,降低对服务器的依赖。RIA 界面具有灵活多样的界面控制元素,能很好地与数据模型相结合,使服务器响应对整个界面产生的影响,转变为只有收到请求的 RIA 应用程序才做出相应变化<sup>[16]</sup>。本系统采用 Flash 作为 RIA 客户端开发平台。

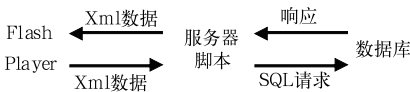


图4 邻近测站观测数据显示处理  
Fig.4 The processing flow chart showing observation data of surrounding stations

#### 2.3.1 地图实现

利用地理信息数据先生成江西省地图的底图和对应的河流、铁路、省道、国道等图形,并导入 flash,制作江西地图的影片剪辑(MovieClip),将底图、水系、县界、铁路、省道、国道等各定义为一个图层,利用 action script 2(Macromedia Flash 脚本语言)编程实现各图层的叠加,以及放大、缩小和移动等基本的地图功能。

#### 2.3.2 台站客户端与服务器端通信

气象台站用户端 Flash 浏览器与 Xml 数据之间以字符串形式直接进行数据交换,flash 调用相关函数通过 http 协议将请求发送到 .NET 服务器,服务器端脚本(本系统采用 asp.net)作为 Xml 数据和数据库的中介,负责将数据库中相关测站观测数据动态生成 Xml 数据并返回台站客户端。

#### 2.3.3 用户端观测数据显示

采用 action script 2 编程实现客户端 flash 接收

Xml 数据,该 Xml 数据包含各邻近站点经纬度、站名、站号和要素测值等,经 Flash 加载解析后,将其中的经纬度转换成地图上对应的坐标,并在地图上显示发生可疑数据的测站及邻近测站的观测数据。

## 3 应用效果

2009年9月至2010年6月,应用本系统对全省87个国家站及1531区域站正点观测数据进行了实时质量控制,共发现疑误数据19236个,其中气压数据8120个(占42.8%)、温度数据8957个(占47.2%)、降水数据765个(占4.0%)、其他要素数据1394个(占6.0%);系统共剔除错误数据14910个,修订可疑数据4056个(台站未能核查修订的可疑数据270个),有效降低了自动气象站疑误数据对气象业务服务可能造成的不利影响。

本系统的业务应用形成了省、市、县三级联动的数据质量实时控制业务流程。台站业务人员能够实时获知、检查分析、修订反馈本地出现的疑误数据,相应市局业务人员可及时掌握并指导所辖台站的疑误数据处理。省级中心对全省自动站资料的自动实时预审时效与资料传输收集时效基本同步,完全可以满足资料分发和应用的实时性要求。可疑数据的检查结果反馈时效则视测站性质的不同有所差异,国家站可疑数据处理的反馈时效一般在1小时内,区域站可疑数据的反馈时效一般在数小时内。

## 4 小结

(1) 本系统将数据质量控制技术和网络消息即时通信技术相结合,以业务流程为主线,构建了省、市、县三级数据质量实时控制业务协作平台,不仅提高了自动站疑误数据处理的效率,还对台站的数据质量意识和数据处理能力的提高起到了较好作用。

(2) 错误观测数据的自动剔除与可疑数据处理的快速反馈,为监测预报等实时性强的业务尽可能提供了准确、及时的观测数据,促进了上述业务服务工作的开展。

(3) 本款业务软件的开发思路也可以在探空、农气、大气成分等观测资料的实时质量控制中参考使用。

(4) 自动站数据质量实时预审所涉及的审核规则与疑误数据的判别结果紧密相关,需要不断地进行审核规则与指标体系的完善。

## 参考文献

- [1] 李伟,陈庆华,王经业.气象探测数据质量问题分析[J].气象软科学,2007,4:77-81.
- [2] 任芝花,熊安元.地面自动站观测资料三级质量控制业务系统的研制[J].气象,2007,33(1):19-24.
- [3] 任芝花,赵平,张强,等.适用于全国自动站小时降水资料的质量控制方法[J].气象,2010,36(7):123-132.
- [4] 王海军,杨志彪,杨代才,等.自动气象站实时资料自动质量控制方法及其应用[J].气象,2007,33(10):102-109.
- [5] 窦以文,屈玉贵,陶士伟,等.北京自动气象站实时数据质量控制应用[J].气象,2008,34(8):77-81.
- [6] 王建庄,许沛林,彭惠英,等.Ⅱ型自动气象站数据采集的实时质量控制[J].广东气象,2009,31(5):57-58.
- [7] 刘小宁,任芝花.地面气象资料质量控制方法研究概述[J].气象科技,2005,33(3):199-203.
- [8] 熊安元.北欧气象观测资料的质量控制[J].气象科技,2003,31(5):314-320.
- [9] 王新华,罗四维,刘小宁,等.国家级地面自动站 A 文件质量控制方法及软件开发[J].气象,2006,32(3):107-109.
- [10] 何志军,封秀燕,何利德,等.气象观测资料的四方位空间一致性检验[J].气象,2010,36(5):118-122.
- [11] 刘小宁,鞠晓慧,范邵华.空间回归检验方法在气象资料质量检验中的应用[J].应用气象学报,2006,17(1):37-43.
- [12] 陶士伟,徐枝芳.加密自动站资料质量保障体系分析[J].气象,2007,33(2):34-41.
- [13] Evjen B, Hollis B 著. VB. NET 高级编程(第三版)[M]. 杨浩译.北京:清华大学出版社,2005.
- [14] [美]Freeman A, Jones A. NET 安全编程(C#/VB. NET)[M]. 靳京译.北京:清华大学出版社,2007.
- [15] 顾宁,刘家茂,柴晓路. Web Services 原理与研发实践[M].北京:机械工业出版社,2006.
- [16] 颜金杪. Flash MX 2004 ActionScript 2.0 与 RIA 应用程序开发[M].北京:电子工业出版社,2005.



## 特色农业气象服务技术与应用

濮梅娟 主编

该书概述了江苏特色农业、农业气候资源、气象灾害情况;阐述了园艺作物、设施蔬菜、水生植物、大宗农作物不同生育期的气象灾害防御措施、病虫害防治技术、农事建议和田间管理措施;介绍了水产、畜禽生长发育的环境条件和养殖技术。该书可供党政部门领导及气象、农业、林业、国土、民政、地理、环保等部门的科技工作者参考应用,也可作为有关院校师生教学参考和借鉴。

16 开 定价: 45.00 元

## 环境科学概论

郑有飞等 编著

该书主要介绍了环境科学的基本概念、问题、特征、原理、方法和技术。全书共十二章。第一章介绍环境和环境科学的定义、发展历程和学科分支;第二章介绍当前人类面临的主要环境问题和环境保护;第三章介绍环境保护中的生态学知识;第四章介绍大气层、大气环境现状和大气的污染防治;第五章介绍水资源和水污染现状、水污染防治方法和原理;第六章介绍土壤环境及其污染现状、土壤污染防治和土壤修复;第七章介绍固体废物产生现状及危害、固体废物处理原则及管理、固体废物治理与资源化;第八章介绍各种物理性污染及其防治;第九章介绍全球环境问题和气候变化;第十章介绍人口、资源与环境的关联和协调发展;第十一章介绍可持续发展战略基本理论及其实施方法;第十二章介绍清洁生产概念、理论、末端治理及全过程控制。

该书可作为高等学校环境类和非环境类专业本科生产专业课和公共选修课的教材,也可作为教师和相关领域广大环保、气象科技工作者的参考用书。

16 开 定价: 40.00 元

## 勤勤恳恳 教书育人——纪念著名气象学家仇永炎教授

北京大学物理学院大气与海洋科学系 编著

该书收录了三部分内容:第一部分是纪念文章,从不同侧面展现了一个普普通通的教育工作者辛勤耕耘、感人至深的一生。第二部分选出了仇永炎教授在气象学教学、中期天气预报、臭氧气候等方面的代表作。第三部分介绍了仇永炎先生在气象学教学和中期天气过程研究等方面的成绩,并汇集了仇永炎先生的弟子及友人在灾害天气、动力气象学、中期天气预报、东亚季风等方面的研究成果。

该书对关心仇永炎教授气象学教学成就的人们,以及对从事天气、气候、台风、寒潮、灾害、大气臭氧、中期预报、大气物理学等方面工作的人员有重要参考价值。

16 开 定价: 88.00 元

## 《气象知识》三十年文萃丛书

《气象知识》编辑部 编

《气象知识》创刊三十年来,刊载了大量的优秀气象科普作品。编辑部辑成《气象知识》三十年文萃(1981—2010)丛书。丛书包括四个分册:《气象灾害面面观》、《万千气象多纷呈》、《晴阴冷暖总关情》和《气候变化纵横谈》。

16 开 定价: 29.00 元