Quality Control and Analysis of Surface Meteorological Station Data

Kewen Xu Guixiang Zhao

(**National Climate Observation Station of Taiyuan,**
No568,pingyang Road, Taiyuan 030006 , Shanxi ,P.R.China)
E-mail:xu411-wang921@163.com

Abstract:

Using conventional observations of automatic and manual weather station from January 2003 to September 2004 (temperature, pressure, relative humidity and soil temperature from 0 cm to 320 cm),a contrast analysis was gave. The results show that: observational errors of temperature, pressure and soil temperature from 40cm to 320 cm are relatively small, while relative humidity, soil temperatures from 5 cm to 20 cm and ground temperature are greater. There are many factors that caused data differences between automatic and manual observation, including different measure principles and observation method, different measure time level, different environment, and so on. It can provide basis for continued use of data observed by automatic station. A real meteorological data quality control is based on right way and fit method.

Key words: automatic weather station, manual weather station, comparative analysis, surface meteorological observation method

1 Introduction

Data are collected on a continuous and long-term basis. Data collection could be done simply by manually entering the forecast and observation values in verification forms that are then sent to the data processing site. The data could be electronically transmitted after being entered by the forecaster or the observer. In the case of scoring systems, collection is easier and should be done much more frequently, and in a highly automatic data collection system the data are retrieved and transmitted to appropriate destinations rapidly, but there are inherent risks of neglect of good quality control. Without reliable data, prediction has no foundation. Observation data should be collected on a regular basis with reasonable imposed deadlines. For highly automatic systems this could be done moment after observation times, but for a system that uses verification forms, collection should be made a few times monthly. All data should be checked prior to calculation of verification statistics. A simple quality control method is that the forecaster check and manually corrects the data. Automatic error-checking is more complex though not necessarily better since small errors can go easily undetected. Sometimes too, if quality control measures are too strict, automatic procedures may eliminate highly unseasonable or rare and extreme values that are correct. These data come from some of the most critical forecast situations and special care should be made to preserve them in the verification sample. Quality control is most effective when done at the data source or local level, by forecasters who are most familiar with

daily weather at their specific sites and can better identify errors. In addition to demonstrating trust in the forecaster by having him check his own work, local quality control can add forecaster acceptance of verification by instilling confidence in the accuracy of the results, and by virtue of participation in the scheme. Quality control is also done at the central level so as to identify errors that were missed at the local level, the results being shared with the local level in order to assist in, and also to improve quality control at the source. There are different measure principles and observation method, different measure time, different time level, and so on. Between data of automatic and manual observation Comparing two objective data of basic meteorological elements, such as air pressure, temperature, precipitation, humidity, soil temperature etc. Difference between data observed by automatic station and manual station were discovered. It shows that observation results of automatic station were conducted.

Our station is one of the global observation systems. During the observatory activity, we learned how to read and record measurements with each of the weather instruments that were built. Now, we are using instruments to record automatic data and comparing with manual data together and we are familiar with how to use the automatic weather data, so that we will be able to read, and record the weather Information automatically, access archived data for previous days. Using ground measurement data, we gave this study.

2 data collection and observe processing.

The main goal of surface meteorological observation is to get information about meteorological elements .Description of a real meteorological data is quality control based on right way and fit method. The basic idea is that successive observations are checked in time . If the point are gauging in time, and the observatory value hasn’t changed, which various experiential knowledge of observers. To examine the rate of change for the current observation, comparisons are made with the proceeding and succeeding observation from the same the estimated value will equal the observed value, and the observation will be a process for better using of meteorological data.

Manual observation measure method is necessary to adjust processing. soil temperature should be put in the first step from deep to shallow (soil temperature differs from temperature, it changes slowly, but ground temperature changes rapidly .In the second step we can observer snow cover or rain, cloud cover, visual range, state of the ground and special phenomena, precipitation amount, and atmospheric phenomena .In the third step obtains temperature, humidity, wind speed and direction, pressure, which are made every hour. When the data loss at some fully automatic stations, it will replace of manual thermometer .the observation should be made in as short a time as possible at the exact hour. The standard of time any observations recorded by clock time have been entered. It is reason has differed from real time observations.

3 contrast and analysis of meteorological data

Now, we think that meteorological observation use automatic weather station data

could in stand for manual weather station. Though the method of meteorological observation between automatic weather station and manual weather station is different, one of the factors inevitable to get a result. The situation of that point of observation is very important to analyze and use those data scientifically and more precisely.

The data used in this paper are that: air pressure, temperature, relative humidity and soil temperature at observation stations of Taiyuan city from January 2003 to September 2004 .

Casual departure X_i: Observing element deviation value between automatic observe and manual observe

Average departure X: Average casual departure monthly

$$X = [\Sigma (X_{manual} - X_{automatic})] / n$$

3.1 Air Pressure

A synoptic observer often measures and observes present and past pressure .Our sensor has been compared it with a mercury barometer on site. Figure1, the left stands for average departure of pressure, our reading values are mostly lower and the mean error (during January 2003 to September 2004) is estimated to be about 0.6 hPa,

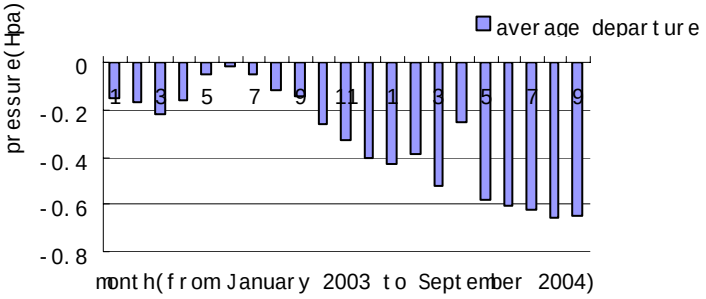

Fig.1: contrasting pressure of automatic weather station and manual weather station

the maximum is about 0.8 hPa ,and the minimum if about 0.1 hPa. This error is gradually increasing, and has been remarkably stable, except that monthly data in April appears to be somewhere greater. It will remain the same value after that.

3.2 Temperature

Figure2, below left shows, is a graph showing automatically temperature and manual surface observations collected during January 2003 to September 2004 ,it always showed consistent values with thermometers and sensor comparative analysis. Temperature readings collected by automatable surface observations are mostly low in winter other than temperature collected by manual surface observations are low in summer .The mean departure is estimated to be about -0.5-0.1℃, with a maximum casual departure of 1.5℃ when temperature changes rapidly. Even if air temperature is consecutive measured at 1 hour apart, finding this error is not easy. For an example ,

September and October 2003 outliers, mean departure reached separately 3.9℃ and 1.17℃ because automatically collecting sensor was break out. It is very difficult to find it. Casual departure is irregularly up and down on occasion. There is little difference between the two stations when it is working normally. In the first step we analyses measurement data characteristics, using methods of correlation and regression, analyses and principal components.

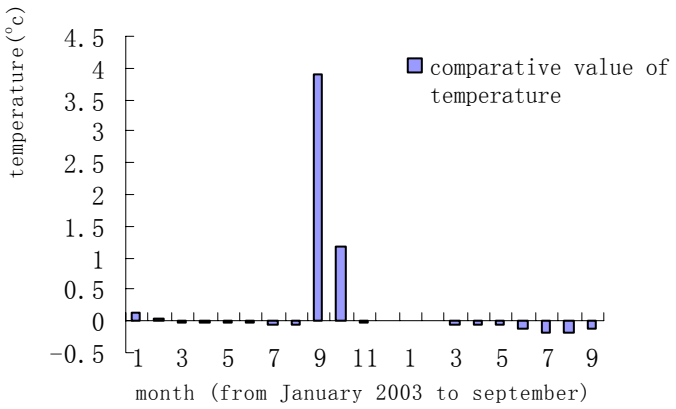

Fig.2:contrasting temperature of automatic weather station and manual weather station

The other test methods are that using analysis of regression and analysis of variance .It shows a choosing diffident 4day for the 24—hour period temperature between -20℃~25℃, consistency checks data between automatic weather station and manual weather station. As an example, we calculated variability of temperature of representative four parts daily, the period of 4 January, 11 January, 14 February, 15 April.

Testing normal distribution sample between automatic weather station and manual weather station , the data regarded has no significant difference on level of significance α=0.05, lining up regression equation and analysis of regression.

Sample test: y shows automatic weather station data

X shows manual weather station data

Variance: S_x^2 expresses automatic weather station data

S_y^2 expresses manual weather station data

For specifying the value, we have to use Fig.3.comparate analysis of variable temperature -20℃～-10℃between automatic weather station and manual weather station .

Hypothesis testing $\delta^2_{manual} = \delta^2_{automatic}$

$S^2_{manual}=11.29$ $S^2_{automatic}=9.99$ $\delta^2_{manual}=10.82$ $\delta^2_{automatic}=9.55$

Sample number: $n1=n_2=24$

level of significance $\alpha=0.05$ $\alpha=0.05$

F distribution, F=1.13

$F_{0.05}(23, 23)=2.01$ $F_{0.95}(23, 23)=0.49$

When $\delta^2_{manual} > \delta^2_{automatic}$,

$F \leqslant F_{1-\alpha}(n_1-1\ n_2-1)$ or $F \geqslant F_{\alpha}(n_1-1\ n_2-1)$ is rejection region

$0.49 \leqslant (F=1.13) \leqslant 2.01$

The examining value of F are greater than $F_{0.05}$, hypothesis is regarded outside rejection region,

$\delta^2_{manual} = \delta^2_{automatic}$,

Two aspects overall amounts have homogeneity of variance

regression equation: y=-0.60+0.93x

Correlation coefficient r=0.994

Using same way made up four runs correlation coefficient and regression equation

table1: Discriminate analysis four section temperature data get correlation coefficient and regression equation of automatic weather station and manual weather station

data / time	S^2_X	S^2_y	δ^2_X	δ^2_y	F	r	regression equation
4 January	11.29	9.99	10.82	9.55	1.13	0.994	y=-0.60+0.93x
11 January	26.83	27.46	25.70	26.32	0.98	0.999	y=0.29+1.01x
14 February	31.02	30.58	29.70	29.27	1.01	0.999	y=0.08+0.99x
15 April	40.83	41.09	39.19	39.14	0.99	0.999	y=-0.02+1.00x

The features are only marginally significant at 95% level on a statistical basis. Base on the comparison analysis between automatic observing and manual observing show that correlation is small difference.It is found that highly significant relationships correlation between automatic and manual and high stability of the run. On the whole, their temperatures are essentially uniform with the environment.

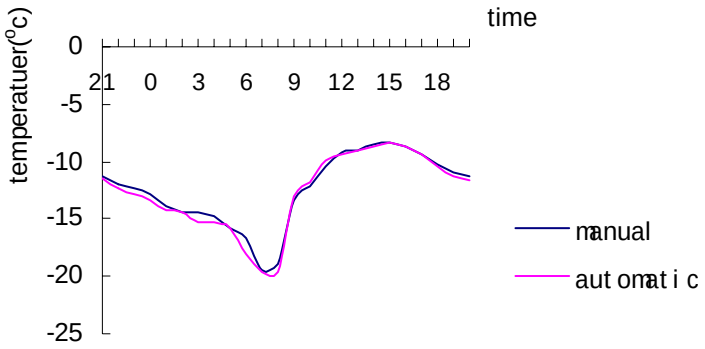

Fig.3:contrasting temperature of automatic weather station and manual weather station(-20°c--10°c)

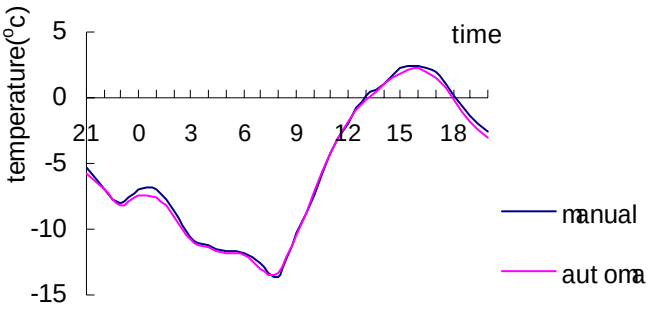

Fig.4:contrasting temperature of automatic weather station and manual weather station (-10°c-0°c)

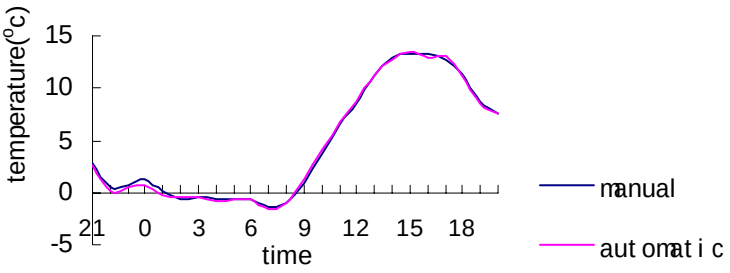

Fig.5:contrasting temperature of automatic weather station and manual weather station(0°c-10°c)

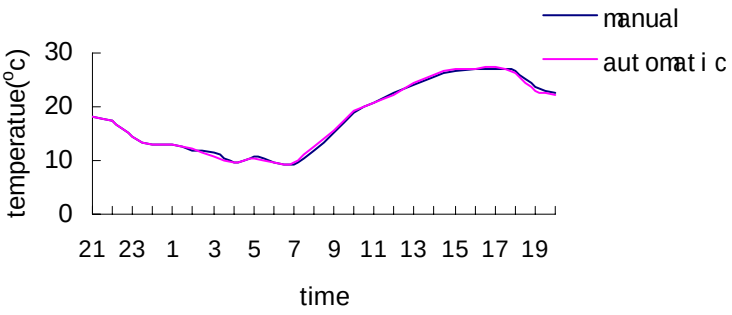

Fig.6:contrasting temperature of automatic weather station and manual weather stationc10°c-25°c

3.3 Humidity

It can be seen from Figure 7 that humidity varies between automatic and manual station from January 2003 to September 2004.There have two high period the monthly average consistent values relative humidity is expressed as percentage varying from March to October, the lowest in November-January .The relative humidity runs instability, especially increasing the influence of spring and autumn. The highest average consistent values is 4 % in April 2003 and June 2003. Hygrometers, when checked during a stable period, proved to be quite reliable; their errors,

usually around 2%-3%, gets bigger for either extremely low or high values. The hygrometer has been substituted，as it was reading about 3% less than very high humidity . The role of the low humidity and the high humidity cases investigate Fig 7.

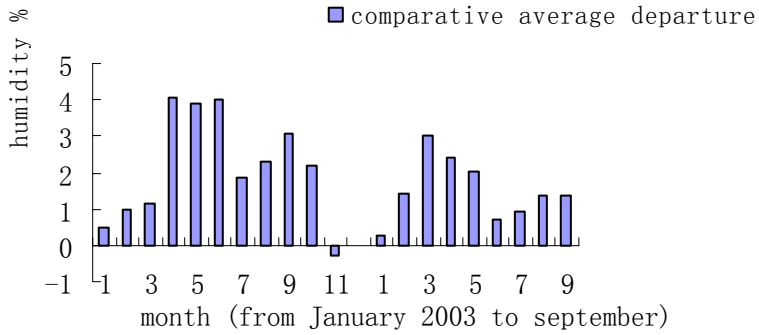

Fig.7:contrasting humidity of automatic and manual weather station

3.4 Soil temperature at a range of depths

Four thermometer sensors were installed above the ground, at ground surface and 5cm deep in the soil. Temperature data are automatically measured with the advance of science and technology, some scientific problems become complicated than they were thought. Sensor is used without the intervention of a human observer, during the experiment; moisture was measured between 0 and 80 cm of the soil. Here, the data shown for past years have been collected from the 1st of January until the 31st of December of a year. Based on the analysis of the characteristics of surface temperature observation, a method of standard comparative for surface temperature observation data was presented 2003.The steps of this method are introduced in detail. The method is essentially the manual checking. In order to gather information regarding current baseline algorithms used by arithmetic departure mean. Data processing, it provides a range of soil temperature at two depths distributions Fig. 8 and Fig. 9.Comparison value is negative number, expressing automatic data is bigger than manual. Otherwise it is the same. Manual data which can differ from one to the other is due to a small difference in the timing of automatic data received. Two depths comparison value are more negative number -1.0℃ from January to September in 2003.The tendency of negative number about 5cm is increasing. When surface has snow covered and using different method ,actually casual departure reaches approximately 10.0℃ sometime. We will also measure the snow thickness at some of these sites.

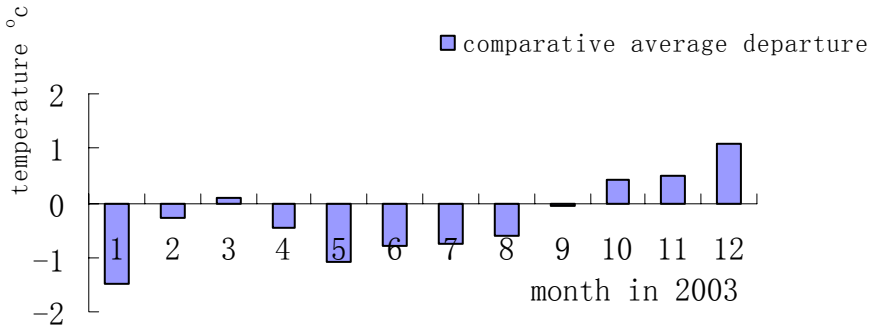

Fig.8:contrasting soil temperature 0cm of automatic and manual weather station

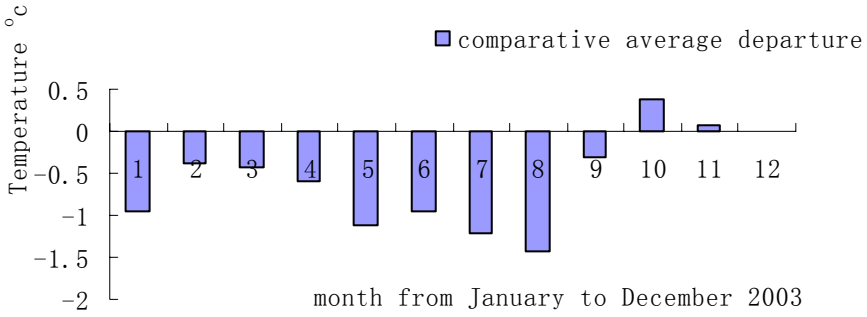

Fig.9:contrasting soil temperature 5cm of automatic and manual weather station

We will use our calibrated numerical models investigate the sensitivity of permafrost temperatures to different scenarios of future climatic change. The first national-level observation is about our experience to reach target of the project. Casual departure of surface temperature is unstable because surface temperature changes rapidly. Surface temperature observation has natural reasons; because the change of sounding environment is different, so using kinds of analyzing methods. The changing of meteorological sounding deep data maintains extremely high stability and seemly deeper more stable.

Advanced

The fact is that automatic station will be used more and more widely. Observatory quality will improve continuously and no observers can be available. The key to weather prediction depends on accurately measurements, these variables and feeding the data. New measurements made by remote sensing instruments using advanced technologies should radically improve weather and climate predictions. Remote sensing is the science of detecting and measuring kind of variables (e.g. temperature and air pressure, humidly and soil temperature) without coming into contact with observing meteorological phenomena of surface observation. (Etc. Light fog and

haze). Meteorological phenomenon has very complex. Meteorological phenomena are in most cases obtained by observers at stations. Until recent years all surface observing have been performed manually; the observer reads the instruments, makes visual estimates of visibility, cloud, present weather and the state of ground, records the observation on paper, this paper was to do a case study based on operational experience.

An additional point the prevailing visibility observe in most cases use traditional methods were discussed .visibility observations made by human observer and must cover a range up to 10 km. We confused that the approach is made somewhere between 0.1km and 10km above the ground like seem phenomena .Typically recently there has been some renewed interest in distinguish between light fogs, smoke and haze. Visibility will be mixed up and use the existing systems by a better way. It is useful to know that visibility observations can be assisted by visibility sensors. We will focus on this point .It is necessary to clearly describe its feature with different conditions.

Observations show that the swapping of sensors can also have a significant effect on the quality of data, frequently introducing discontinuities into a data series. The usefulness of the data obtained from a sensor is heavily dependent on the calibration of the sensor. For data to be comparable with other sites and networks, the calibration of sensors needs to be traceable back to common standards. This is often difficult to establish, particularly with cheaper sensors, but is of equal importance regardless of the quality of the sensor. The quality of the final data received by the researcher can only be as good as the quality of the sensors used. No post analysis of the data can improve the accuracy or reliability of the information obtained.