

## Algorithms for computerised monitoring the performance of DCP network

M. K. GUPTA

*Meteorological Office, Bhuj - 370001, India*

*( Received 13 September 1999, Modified 9 August 2000 )*

सार - भारत मौसम विज्ञान विभाग 1986 से आँकड़ा संग्रहण केन्द्र (डी.सी.पी.) के नाम से सौ स्वचालित मौसम केन्द्रों के संजाल का प्रचालन कर रहा है। इन सभी केन्द्रों में कोई कार्मिक कार्यरत नहीं है, अतः इन केन्द्रों से सम्प्रेषित किए जाने वाले आँकड़ों से प्राप्त सूचना के अलावा आँकड़ा संग्रहण केन्द्र (डी.सी.पी.) के उपकरणों की कार्यप्रणाली को जानने का कोई और साधन नहीं है। अतः आँकड़ा संग्रहण से प्राप्त आँकड़ों का विश्लेषण करते हुए, जो कि उनकी प्रभावी एवं कुशल अनुरक्षण के लिए अनिवार्य है, आँकड़ा संग्रहण केन्द्र के विभिन्न उपतंत्रों की कार्यप्रणाली का मूल्यांकन करने के लिए समुचित एल्गोरिथिम विकसित किए गए। इन एल्गोरिथमों को विकसित करने के लिए जिन अवधारणाओं का उपयोग किया गया है उनके विषय में इस शोधपत्र में चर्चा की गई है।

**ABSTRACT.** India Meteorological Department (IMD) has been operating a network of one hundred Automatic Weather Stations called Data Collection Platform (DCP) since 1986. All these stations are unmanned and there is no way to know the working status of DCP equipment except to extract such information from the data transmitted by them. Hence suitable algorithms were developed to evaluate the working status of various sub-systems of DCP stations by analysing the data received from them, which is essential for their effective and efficient maintenance. The concept used in developing these algorithms is described here.

**Key words** - Data collection platform (DCP), Automatic weather station (AWS).

### 1. Introduction

A DCP station consists of an electronic equipment attached with various electrical sensors for measurement of atmospheric variables such as temperature, pressure, wind speed, wind direction, humidity, rainfall, etc. This equipment is powered by batteries which are normally under float-charge through solar panels. The equipment powers the sensors for a short duration at every full hour, signal conditions the analog signal coming from sensors, converts them into digital data and stores into own memory. It further transmits the stored data three times during the assigned transmission slot in speed-random burst mode toward the communication satellite *i.e.* INSAT. The DCP system operated by IMD has uplink frequency of 402.75 MHz and a burst length of 422 bits. The satellite relays the same data towards earth which are received at INSAT Meteorological Data Processing System (IMDPS), converted into engineering values and disseminated to users.

India Meteorological Department has been operating a network of one hundred such DCP stations as shown in Fig.1 since 1986. All these stations are unmanned and

there is no way to know the status of working or malfunctioning of the DCP equipment or any of its sub-systems except to extract such information from the data transmitted by the stations. The pre-knowledge of the working status is an imperative need for efficient and effective servicing. In the absence of such feedback about functioning of various sub-system, the DCP maintenance personnel, visiting these unmanned stations have to waste their time and energy in testing and calibrating the entire system and may happen that some minor defect may go unnoticed. Therefore, the monitoring of the performance of these stations is essential for proper feedback to maintenance personnel.

The monitoring of the performance of these stations has been a quite tedious and extensive work. The various methods so far adopted using DCP data available over Global Telecommunication System (GTS) network do not reveal the complete picture about functioning of various sub-system of a DCP Station. Moreover, the results are mostly qualitative and do not reveal any picture about improvement or deterioration in the performance. Hence a need was felt to develop a computerised performance monitoring technique.

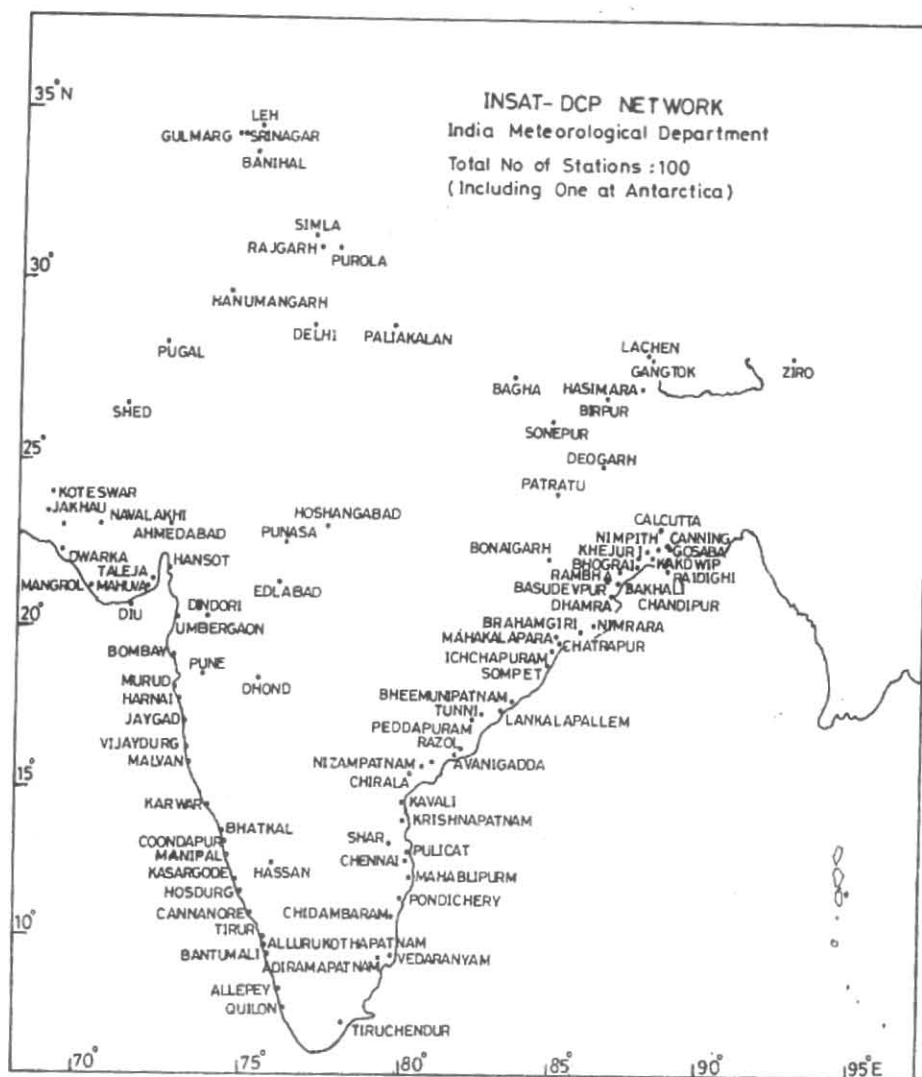


Fig. 1. Network of INSAT-Data collection platforms in India

## 2. Data used and methodology

As described above, the DCP system has several sub-systems *viz.*; R.F. Transmitter, batteries, three test data called calibration voltages and ten sensors. The malfunctioning of one or more of these sub-systems has some unique effect on the data transmitted by the system. All the possible defects associated effects on DCP data have been studied thoroughly and this correlation between system defect and effect on data is used to prepare suitable algorithms to evaluate the status of working or malfunctioning of various sub-systems of the DCP equipment.

The DCP data stream contains some fixed bit fields *viz.*, station and sensors' identification codes. The

accuracy of their reception determines the performance of communication, which in turn helps to evaluate the status of working of R.F. transmitter, batteries and circuits generating fixed bit fields. While the remaining bit fields represent the meteorological data measured at the last full hour. These are variable and are the only source to evaluate the functioning of sensors. A definite limit for each met parameter can be set beyond which they can be considered non-realistic and unreasonable. A reasonable limit can also be set for their variability, (measured here as twice of standard deviation) too. Some of the meteorological parameters have, in addition, a definite diurnal trend like temperature, pressure etc. Thus the accuracy of each meteorological parameter is checked and functioning of associated sensor is assessed.

These algorithms mostly uses arithmetic mean (A.M.) and standard deviation (S.D.) of various data values to check their validity. Therefore, use of one or two days data does not yield a satisfactory result, mainly, because of insufficient quantity of data to calculate representative A.M.'s and S.D.'s. It has been found optimum to use 7 to 15 days raw DCP data to perform the analysis. Further longer duration may not yield a satisfactory result again because of seasonal effect on data.

The raw DCP data received at INSAT Meteorological Data Processing System (IMDPS) is the input to the computer program developed on the basis of these algorithms. The raw data set consists of all the bursts received in random order from all the working DCP stations. All the bursts in this data set are first decoded into corresponding engineering values along with error-status of various fixed bit fields, viz.; station identification, parity, sensor identification, and health bits. This pre-processed data is then separated station-wise for further analysis. The various steps involved in the station-wise analysis are given below.

### 2.1. Communication performance

Three burst per hour are expected from a station but some are severely mutilated and some completely lost in communication channel due to various reasons. The weak R.F. transmitter and weak batteries are the main cause which reduces the transmitted power. Further, some of the received burst despite having errors in few of the bits are not ignored at IMDPS as the station could be assigned and remain useful to some extent. There are 422 bits per burst, of which some are fixed, viz.; station identification, parity, sensor identification, and health bits. The accuracy in reception of these fixed bit fields is used to monitor the communication performance and in turn the status of R.F. transmitter, batteries and circuits generating fixed bit fields is determined.

First, the following two quantities are calculated using the entire data under analysis-

- (i) Percentage of missing burst.
- (ii) Percentage of bursts having bit errors in fixed bit fields.

Since the same data burst is transmitted three times, the data can successfully be retrieved even if 33% of the bursts are correctly received. Therefore, the link performance and hence the working status of R.F. transmitter, batteries and circuits generating fixed bit

fields is taken as satisfactory, if both the above two quantities are less than 30% each. Otherwise the further analysis is done to diagnose the problem as below.

(a) Whether the performance is uniform on all the days during the period. This is necessary to know because the station might have improved/ deteriorated in its performance during the period. This is achieved by determining the above two quantities day-wise and comparing them with the average performance for the period.

(b) Whether the day time performance is better than that at night time. If so, it indicates weak batteries which are not retaining charge properly. This is achieved by determining the above two quantities separately for day-time (0300 UTC to 1400 UTC) and night-time (1500 UTC to 0200 UTC) and inter-comparing them.

(c) Whether bit errors are in few fixed field. If so, it indicates defect in corresponding circuit. If the bit errors are found spread out in almost all the fixed bit fields, there is a possibility of weak R. F. box or weak batteries. For this purpose, percentage of burst having erratic bits in each of the fixed bit field is determined and compared with the average performance. Fixed bit fields having significantly more contribution to erratic burst are counted and reported.

### 2.2. Calibration voltages

There are three calibration voltages named as Cal-I, Cal-II and Cal-III which are 0, 2.5 and 5.0 V DC. The accuracy and stability of these data depends on the status of A/D converter and the circuits generating these voltages. The procedure adopted is simply to find out the percentage of occasions on which these voltages are found out of tolerable limit of  $\pm 150$  mV. The result is reported as such in the output.

### 2.3. Dry bulb temperature (DBT)

This parameter is variable and can only be checked for reasonable variability. But before processing further it is necessary to remove the data which might have got mutilated during communication and had become completely absurd and unrepresentative. For this purpose A.M. and S.D. are determined and data values beyond  $A.M. \pm 3 \times S.D.$  are filtered out for subsequent processing. New A.M. and new S.D. are again calculated and variation range is taken as  $A.M. \pm 2 \times S.D.$  Variability of less than  $4^\circ C$  indicate sensor not responding properly and greater than  $16^\circ C$  indicates abnormal variation. Two separate A.M.'s are calculated for day-time (0300 to

1400 UTC) and night-time (1400 to 0200 UTC) data and difference is calculated to check for proper diurnal variation. The diurnal difference from 4 to 20° C is taken as normal.

#### 2.4. Wet bulb temperature (WBT)

The same procedure is adopted as used for DBT to filter out absurd values and remaining data is used to determine A.M. and S.D. The variation range is taken as  $A.M. \pm 2 S.D.$  and is checked for reasonable variability of about  $\pm 2$  to  $\pm 10^\circ C$ . In case variability found normal, the A.M.'s of DBT and of WBT are also checked for positive difference, *i.e.* DBT - WBT.

#### 2.5. Wind speed

After filtering out absurd data, the remaining data is used to determine A.M. and S.D. Variability of more than 2 knot is taken as normal. But variability of more than 5 knot along with the A.M. value less than 1 knot indicates the sensor not responding at low wind speed.

#### 2.6. Wind direction

Because of circular nature of data, the algebraic A.M. may not be representative. Therefore a special technique involving following steps is used to discard absurd data.

- (i) All values greater than 360 are counted and filtered out. If rejection is more than 20%, there may be calibration error or sensor malfunctioning.
- (ii) The number of data in four quadrant are separately counted and quadrant with maximum count is determined.
- (iii) The A.M. and S.D. of data lying in the most probable quadrant is determined. Twice the S.D. is taken as variability and its value of less than 10 indicates jammed wind vane.

#### 2.7. Pressure

DCP pressure sensor is of mechanical type and suffers from contact problem in the micro-potentiometer. The open contact may cause reporting of Datum pressure or datum+100hPa depending upon type of DCP equipment in use. Therefore, the first step in filtering is to count separately for values less than datum+2 hPa and greater than datum+98 hPa and to filter out all such data. If any of these counts are more than 20% of the total burst received, the error is reported. A.M. and S.D. of

filtered data are determined and further filtering is done taking the acceptable range as  $A.M. \pm 3 \times S.D.$  Again new A.M. and S.D. are determined and variation range is taken as  $A.M. \pm 2 \times S.D.$  If variability is less than 1 hPa or greater than 10 hPa, the pressure sensor is reported as not working satisfactorily.

Further, the approximate lower and upper limit of SLP are taken as  $(990 - \text{altitude}/10)$  and  $(1025 - \text{altitude}/10)$  respectively. Percentage of data separately below the lower limit and above the upper limit are determined and reported if either is greater than 10%. Further analysis is performed to check for proper semidiurnal variation. Four averages are determined for data belonging to 2200-2300, 0400-0500, 1000-1100 and 1600-1700 UTC durations and these are inter-compared and result is reported.

#### 2.8. Humidity

Data is first filtered out as for DBT, A.M. and S.D. of the filtered data is determined. If A.M. is less than 10 or variability less than 10, it is reported as sensor not responding properly.

#### 2.9. Rainfall

DCP system reports total rainfall amount since its switching ON or resetting of the rainfall counter circuit. The reported count should either remain constant (no rainfall) or increase monotonously (during rainfall). All values which do not observe this criteria are filtered out. In case rejected data is more than 20%, it is reported as erratic working of R.F. counter. If the remaining data have same value, it is reported as no R.F. or sensor not working. If the reported value is always zero, it is reported as the R.F. signal conditioner card is not working.

### 3. Product and results

A software package has been developed using the above methodology and performance reports for various stations have been generated using raw DCP data for various periods. Two such report for station Pune and Hassan for the period 20th to 30th December 1997 are at Appendix A and B for stations working satisfactorily and unsatisfactorily, respectively. The algorithms have been found successful in predicting the likely defects at the DCP stations. The performance reports of all working stations for consecutive periods were also compared and found consistent.

### 4. Conclusion

These are quality control algorithms which uses various limit to declare the performance as satisfactory or

not. At this moment certain limit have been optimised on the basis of presently working stations. Some of these limits may need optimisation in due course of time. Also the limits may be changed to tighten the quality control.

These algorithms provide quantitative measures of performance for various DCP sub-systems. This will help in keeping the statistics of performance of the DCP stations and their sub-systems. In addition to above, the results can be used to declare a station reportable or non-reportable for data dissemination purpose to avoid unreliable data reaching the user.

*Acknowledgements*

The author is very grateful to Shri Rudrapratap for his invaluable assistance in running and testing the software during its development. Thanks are also due to Shri P.L. Jamkhande for providing assistance in typing.

**Appendix - A**

<i>DCP performance report for duration</i>	: 20-30 December 1997
<i>DCP Station</i>	: Pune (301)
Total bursts expected during the period	: 720
Total bursts received	: 517
Percentage of missing bursts	: 27 %
Percentage of bursts having bit errors in fixed bit fields	: 18 %
Communication performance	: Satisfactory
<i>CAL-I ( 0 volt )</i>	
No. of bursts with cal-I > 0.2 V	: 0.1%
<i>CAL-II ( 2.50 Volt )</i>	
No. of bursts with 2.3 > cal-II > 2.7 V	: 0.3%
<i>CAL-III ( 5.00 volt )</i>	
No. of bursts with cal-III < 4.8 V	: 3.8 %
<i>Dry bulb temperature</i>	
No. of bursts with abnormal values	: 2.7 %

Variation range	: 18.8±12.6 ° C
Variation found normal. Sensor seems working	
Diurnal variation	: NORMAL (10.2 ° C)
<i>Wet bulb temperature</i>	
No. of bursts with abnormal values	: 0.5 %
Variation range	: 0.1 ± 0.2 ° C
Variation too small. Sensor seems not working	
<i>Wind speed</i>	
No. of bursts with abnormal values	: 0.3 %
Variation range	: 1.0 ± 3.8 knot
Variation found normal. Sensor seems working	
<i>Wind direction</i>	
No. of bursts with abnormal values	: 1.1 %
Variation range	: 205 ± 82 Deg.
Variation found normal. Sensor seems working	
<i>Station level pressure</i>	
No. of bursts with abnormal values	: 0.5 %
Variation range	: 951.2±2.4hPa
Variation found normal. Sensor seems working	
Semi - diurnal variation	: NORMAL (2.8 hPa )
<i>Relative Humidity</i>	
No. of bursts with abnormal values	: 0.5 %
Variation range	: 77.5 ± 39.7 %
Variation normal. Sensor seems working	

**Appendix - B**

*DCP performance report for duration* : 20-30  
December 1997

*DCP Station* : Hassan (300)

Total bursts expected during the period : 720

Total bursts received : 345

Percentage of missing bursts : 51 %

Percentage of bursts having bit errors  
in fixed bit fields : 41 %

Communication performance : Unsatisfactory  
Analysis for unsatisfactory  
communication

\* Performance significantly unsatisfactory on date(s)

Date	Total bursts received	Percentage good
	----- None -----	

\* No significant difference in Day / Night time  
performance

\* Bit errors found frequently in -

Location	Occurrence of error
* Sensor ID of wind direction	: 13 %
* Sensor ID of pressure	: 14 %
* Parity of rel. hum.	: 10 %
* Sensor ID of rel. hum.	: 15 %
* Parity of rain fall	: 14 %
* Sensor ID of Rainfall	: 18 %
* Parity of rainfall overflow	: 18 %
* Sensor ID of rainfall overflow	: 20 %
* Parity of sunshine duration	: 22 %
* Sensor ID of sunshine duration	: 24 %
* High health bit No.3	: 93 %
* High health bit No.4	: 93 %

Above indicates possibility of weak R.F. Box

*CAL - I ( 0 volt )*  
No. of bursts with cal-I > 0.2 V : 1.7 %

*CAL-II ( 2.50 volt )*  
No. of bursts with 2.3 > cal-II > 2.7 V : 4.9 %

*CAL-III ( 5.00 volt )*  
No. of bursts with cal-III < 4.8 V : 7.2 %

*Dry bulb temperature*

No. of bursts with abnormal values : 4.0 %  
Variation range :  $20.1 \pm 5.7^\circ \text{C}$   
Variation found normal. Sensor  
seems working  
Diurnal variation : NORMAL  
( $4.1^\circ \text{C}$ )

*Wet bulb temperature*

No. of bursts with abnormal values : 0.8 %  
Variation range :  $0.0 \pm 0.9^\circ \text{C}$   
Variation too small - Sensor  
seem not working

*Wind speed*

No. of bursts with abnormal values : 0.5 %  
Variation range :  $0.1 \pm 0.6$  knot  
Variation too small - Sensor  
seems not working

*Wind direction*

No. of bursts with abnormal values : 92.4 %  
Sensor seems defective - Mostly  
reporting erratically.  
No. of bursts with WD = 0 : 92 %  
10 k Pot seems defective (open).

*Station level pressure*

No. of bursts with abnormal values : 13.6 %  
Variation range :  $908.5 \pm 1.7$  hPa  
Variation found normal. Sensor  
seems working  
Semi - diurnal variation : NORMAL  
(1.6 hPa)

*Relative humidity*

No. of bursts with abnormal values : 12.7 %  
Variation range :  $50.3 \pm 19.8$  %  
Variation normal. Sensor