Indian J. Met. Hydrol. Geophys. (1975) **26**, 3, 302-308

551·501·7

Automatic processing of synoptic and upper air data

R. P. SAXENA, T. K. MUKERJI, Z. E. SHAIKH and R. K. DATTA

Meteorological Office, New Delhi

ABSTRACT. All meteorological data are processed by a computer to extract Synop and upper air data and to decode them. The decoded data are stored in the memory in a manner suitable for numerical analysis and prognosis. This processing by computer also checks erroneous values of data received or observations not received and gives suitable printed information regarding them.

In addition to the extraction of data, a programme of hydrostatic check has also been incorporated to check the accuracy of the height and temperature values reported.

1. Introduction

Meteorological data, in coded form, becomes available in the forecasting offices within a short period of the time of observations. This is manually decoded and plotted on weather charts for further analysis. This is a time consuming process and is not very efficient, when the volume of data is large. In addition, for computerised analysis and prognosis, it is desirable to transfer the data directly from the teleprinter tapes to the memory of the computer, so as to eliminate the intermediate process of punching the data in the punch cards. As such, the necessity of automatic processing of meteorological data arises and is tackled in this report in a simplified manner. As a first step, a programme has been written, which is capable of reading the incoming data in the coded form and distinguish the Synop, Pilot and Temp. data pertaining to a given period of observation. It isolates these data and then goes on to decode them. Any departure from the standard weather codes is detected and messages written out to indicate the same. After the data sequence is over, the accepted data is printed out followed by the index number of stations, whose data has been rejected. For the Temp. data alone, a programme of hydrostatic check is next carried out and the corrected data is stored and printed out in a proper format.

2. Weather codes utilised

The Synop, Pilot and Temp. codes selected for this study are given in Appendix I. They pertain to 00 GMT and 12 GMT data from land stations.

The Temp. data for the blocks Nos. 50 to 59 is received in two parts, one containing data up to 500 mb and other for 400 to 100 mb. The relevant code is available in WMO Publication F.M. 35 C code.

3. Library of station index number

The stations from which observation are to be received and decoded are selected in advance. The block-numbers and index-numbers of these stations are then arranged according to whether they report wind speed in km/hr or in knots. Thus it becomes known, how many stations are of each type and block/index number checking is made according to that. This will be explained further in the discussion of subroutine DTM. These arranged block numbers and index numbers are read into the computer, thereby creating a library of stations. This library can be expanded, if additional reporting stations are added. At present, we have provision of 350 stations.

4. Selection of Synop, Pilot and Temp. data

The logic of selection of the Synop, Pilot and Temp. data has been shown in the simplified block diagram in Fig. 1. The first line of incoming data is read by the computer. It now checks if the first four digits are LAST or ZCZC. When it is LAST, it indicates the end of data. In the normal coded messages this indicator is not available and has been introduced here to end the programme. Normally an equality sign is transmitted at the end of data of a station. If the first four digits are ZCZC it indicates the beginning of messages and as such, the next line of data is now read. The first two digits of next line are indicators and afford a branching to different types of data. Thus if they are US, it indicates that the message following will contain Temp., if the first two digits are SM or SI, then Synop data will follow, and if the first two digits are UI, the message following will be Pilot. The distinction can also be made to separate out, if the data belongs to block 50-59 (China data)* for which the branching is made to China routine. On the other hand these two indi ator digits are UE,

* Now being received in the International Code.

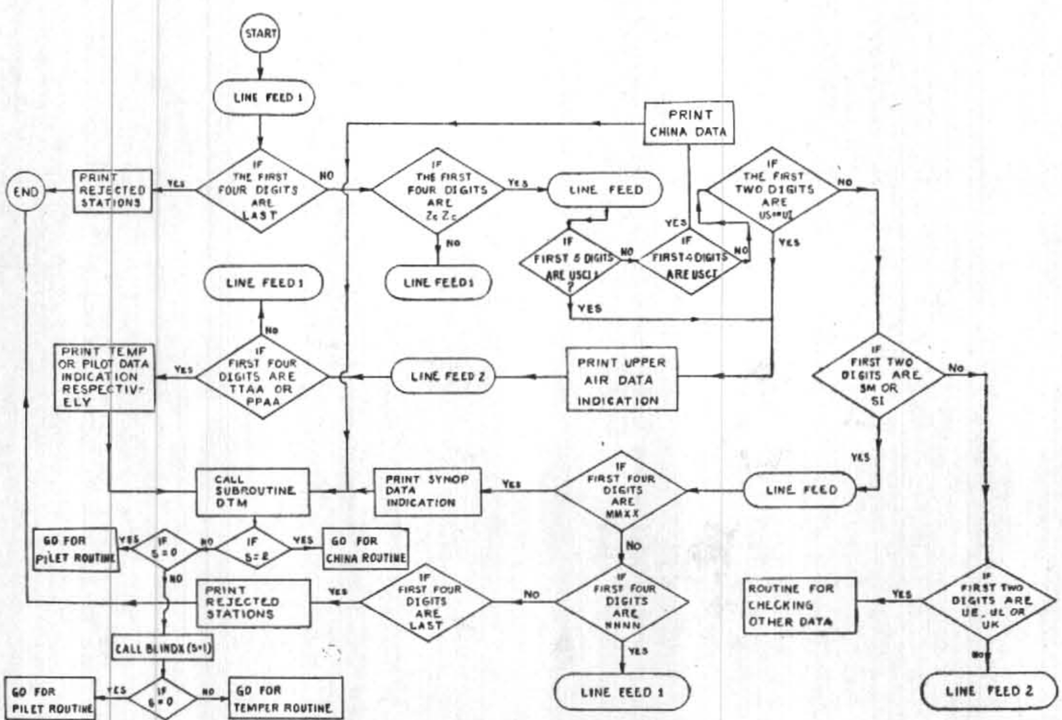


Fig. 1. Block diagram for selection of SYNOP, PILOT and TEMP data

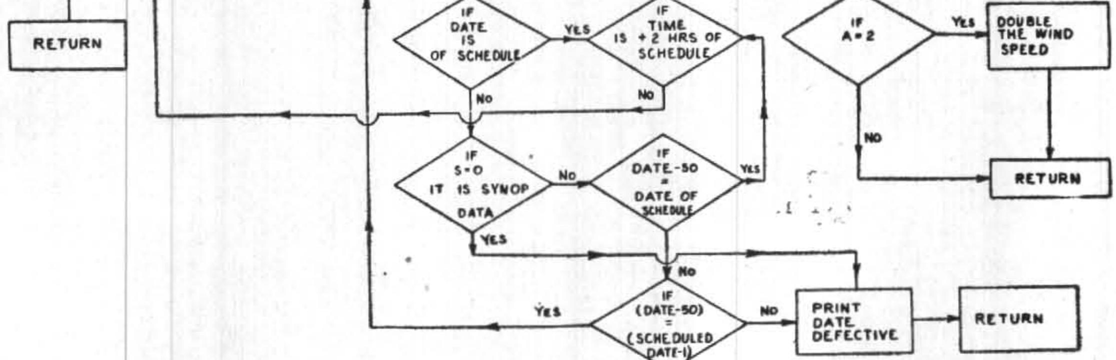


Fig. 2. Diagram for checking date and time : Subroutine DTM

UK, UL, UU etc. and in that case different routines for further checking could be called in. Once the type of message is determined, more data is read and checks are made for the date and time of the message.

5. Subroutine for checking of date and time

Before the data starts to flow in, the computer is supplied with an indicator card specifying the date and time of the message to be extracted. In the checks for date and time (Subroutine DTM), all the incoming messages are scanned with reference to this specified date and time and only those are selected which meet the laid down criteria for these.

In this subroutine, apart from checking the date and time, in case of Temp. data, we also check if the wind speed has to be doubled. This is necessary because, many stations report the wind in km/hr and others in knots. This fact is indicated in the message by adding 50 to the date group, if the reporting is in knots. So, for uniformity the wind speed when reported in km/hr is doubled to convert it into knots.

Time and date reported in a message is considered correct in case of Synop messages if it is same as the specified date and time. However in Temp. messages, provision has to be made for the cases where the 00 GMT observation may be reported as 22 GMT observation of previous date, or 02 GMT observation of the same day. For this reason previous day's data is considered correct if the time reported is within —2 hr of the schedule. Time reporting, similarly allows for + 2 hr on the same day, as being correct. In case, the date or time group falls outside these limits, the data is rejected and a message to that effect is printed out.

6. Subroutine for checking of station index number

Once the computer reads the five digits giving the station block and index numbers, it has to decide if that station data should be accepted or rejected. For this the following criteria are laid down —

(*i*) The station index number should be available in the library of station indices mentioned before.

(*ii*) Same station data should not be repeated. In case of repetition, the repeated data will be rejected. However, in case of China data, duplicate stations are permissible because of split message transmitted for these blocks.

(*iii*) In case of Synop data, further checks have to be made to determine if the wind speed reported requires doubling.

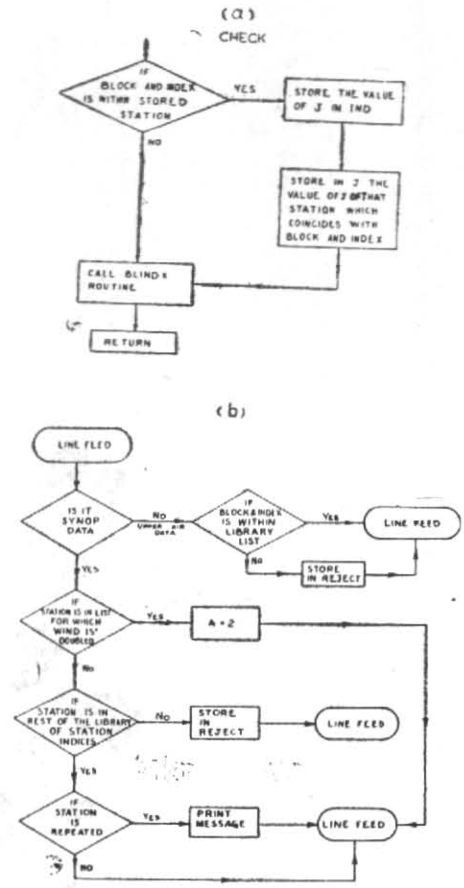
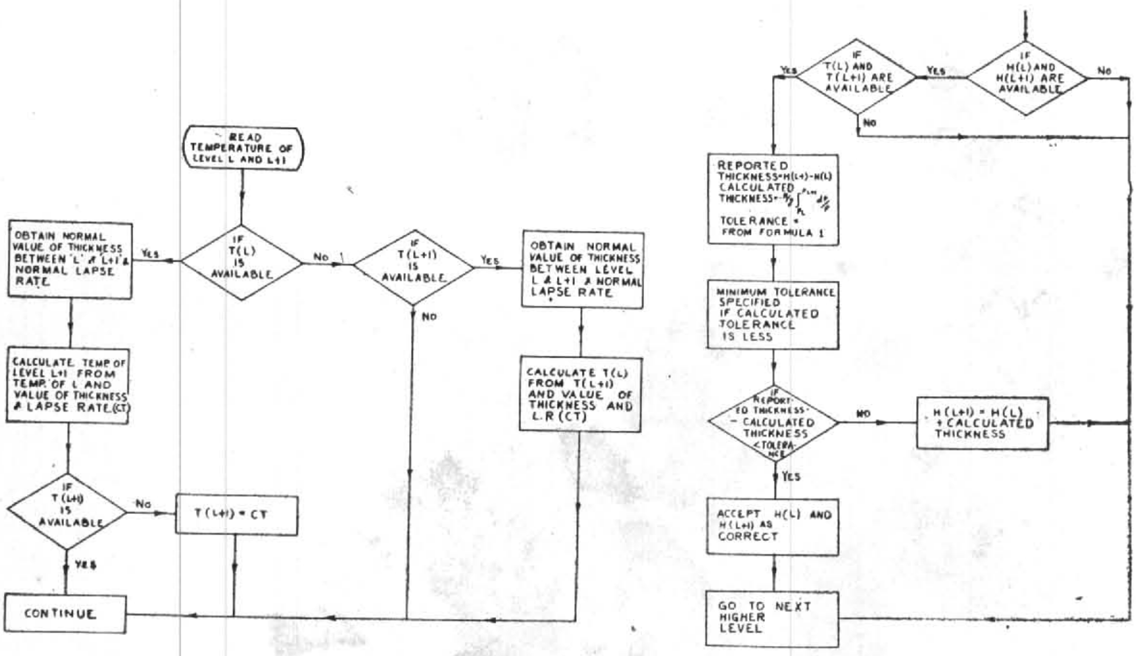


Fig. 3. Block diagrams for (a) checking duplicate station index number and (b) checking of station index number : Subroutine BLINDX

First a decision is made as to the nature of the data whether it is Synop or Temp. In the latter case, the check for block and index is done straightaway with the library of station indices. In case it is not in the list, the index number is stored under rejects and printed out later. Otherwise, it proceeds to read more data. However, in case the data pertains to Synop, the station index number is checked with those stations in the library, for which the wind speeds are to be doubled. If it is not within that group, it is compared with the rest of stations in the library. If it is in that set, more data is read, if not it is stored under rejected stations. In between a further check is made for ascertaining whether the station data has been repeated. If it is repeated, a message is printed out. If it is not, then the decoding of the data continues further.

7. Decoding of the Synop message

As can be seen from the Synop code, given in Appendix I the number of weather phenomena reported in these messages vary with time of observation and the type of reporting station. There are quite a few optional groups



(a) (b)

Fig. 4. Block diagrams of hydrostatic check routine for (a) Temperatures of different levels and for (b) Different pressure levels

For this study, only the 00 GMT and 12 GMT reports from land observatories have been considered. Also, at present, we have provision to check only the mandatory groups.

The decoding is done simply by reading the different groups and storing them under proper headings, after suitable modifications. Thus wind direction is stored after multiplication by 10 and wind speed is multiplied by 2, if necessary. Pressure groups are converted into values which are ready for plotting. Temperatures and pressure changes are checked and converted to proper negative values, whenever necessary. All the decoded values are stored in the memory of the computer against the respective index numbers. A print out of the data is also provided before the next station's data is accepted for decoding.

It will be clear from the above procedure that the decoding of Synop data from ships, hill stations, coastal stations, etc. can be achieved in a similar manner. Since, ship reports are preceded by their coded names and not any index numbers, a library of the same will also have to be prepared.

8. Decoding of Temp. messages

Once a Temp. message is sensed, a number of different subroutines are brought into action to decode different portions of the message and store the same. Thus one subroutine is utilised to decode the temperature, dew point, wind direction and wind speed for different isobaric levles. Once these values are available, another subroutine is called in to store these values. In addition, the heights of different pressure levels are also computed in this subroutine.

In the first subroutine the temperatures are converted into negative values where necessary and wind speed is doubled if they are reported in km/hr. If the wind direction is reported to be greater than 360°, due to oversight or mutilation, a message is printed out and direction is stored as 999. We thereby assume that the wind speed cannot be greater than 300 kt.

The second subroutine is utilised to compute the heights of standard isobaric levels, convert them into heights ready for plotting and store them. Since the same subroutine decodes the heights of 850, 700, 500, 300 and 200 mb, a check is applied to find the level first. Depending on that, the conversion is made so as to give heights correctly. This process is continued till a group LAST is encountered. In that case, the hydrostatic check routine is called in. After that check is over, the rejected station indices are printed out. The complete set of data is then written on magnetic tape in a suitable format ready for machine analysis.

The provision has also been made to merge first part with the second part of the China data under the same station. Besides, the programme also decodes the tropopause and maximum wind data.

The list of various subroutines in the package are given in Appendix II.

9. Hydrostatic check

In this subroutine, the temperatures and height values for different pressure levels, obtained after decoding in the previous parts of this programme, are subjected to checks to ascertain their correctness. This can be broken up into two parts :

1. Checking of temperatures, and
2. Checking of contour values.

The flow diagram for the checking of temperature is given in Fig. 4(a).

The temperature of a level $T(L)$ is read first. Since in the decoding process we have replaced the temperature by 999 in case it is mutilated, its availability or otherwise is checked at first stage. If it is available it is accepted as correct, only if it is within 2°C of previous day's temperature. In case $T(L)$ is not available $T(L+1)$ is read and its value checked. If $T(L+1)$ is correct, $T(L)$ is calculated from $T(L+1)$ with average value of lapse rate and average value of thickness between levels L and $L+1$. On the other hand, if $T(L)$ is available but $T(L+1)$ is not available, the value at $T(L+1)$ is computed from that of $T(L)$. If both $T(L)$ and $T(L+1)$ are not available, then the next level's temperature is read and the process continued. Since for the checking of height values of constant pressure levels, it is necessary to have the temperatures at all the levels. All the reported temperatures are corrected and the missing ones are computed.

The next part of the programme checks the reported contour values. The flow diagram is given in Fig. 4(b). Once heights of level L and $L+1$ given by $H(L)$ and $H(L+1)$ are read, corresponding temperatures are also checked for availability. In case all these are available the computation for thickness and tolerance proceed as follows. Reported thickness is given by $H(L+1)$—$H(L)$. The calculated thickness is obtained by integration of $dp = -g\rho dz$.

Tolerance is computed in the following manner. DALR is assumed from level L to $L+1$ and then from level $L+1$ to L (Fig. 5).

Assuming $T = T_L\,(P/P_L)^k$

$$H_a = -\frac{R}{g}\int_{P_L}^{P_{L+1}} T \cdot \frac{dp}{P}$$

$$= \frac{R}{gk}\, T_L\,[\,1 - (P_{L+1}/P_L)^{k}\,]$$

Next $T = T_{L+1}\,(P/P_{L+1})^k$, then

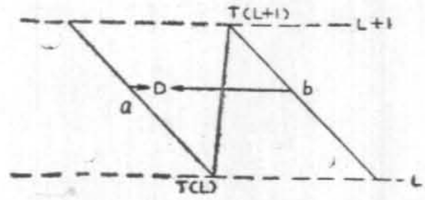


Fig. 5. Computation of height tolerance

$$H_b = \frac{R}{g}\int_{P_{L+1}}^{P_L} T_L \left(\frac{P}{P_{L+1}}\right)^k \frac{dp}{P}$$

$$= -\frac{R}{gk}\; T_{L+1}\left[1 - \left(\frac{P_L}{P_{L+1}}\right)^k\right]$$

Then, Tolerance $= \frac{3}{4}$ of $\frac{1}{2}\,|\,H_a - H_b\,|$

$$= \tfrac{3}{8}\cdot \frac{R}{gk}\left[\,\left|\,T_L - T_L\left(\frac{P_{L+1}}{P_L}\right)^k + T_{L+1}\right.\right.$$

$$\left.\left. T_{L+1}\left(\frac{P_L}{P_{L+1}}\right)^k\,\right|\,\right]$$

$$\text{or, Tolerance} = \tfrac{3}{8}\left(\frac{R}{gk}\right)\left|\;\left[\,1 - \left(\frac{P_{L+1}}{P_L}\right)^k\,\right]\right.$$

$$\left.\left[\left(\frac{P_L}{P_{L+1}}\right)^k T_{L+1} - T_L\,\right]\right|$$

The value of tolerance will change with the lapse rate. In case it is too little, it is increased to 20 gpm upto 500 mb, 40 gpm upto 200 mb and 80 gpm for layers above that.

After all these are calculated, the difference between reported thickness and calculated thickness is compared with tolerance. If it is within tolerance, it is accepted, if not, the value of $H(L+1)$ is put equal to $H(L)$ *plus* calculated thickness. The process is then repeated. It may be noted that if any of the quantities $H(L)$ $H(L+1)$, $T(L)$ and $T(L+1)$ is not available the process cannot work without going to next higher level data. However, this difficulty will occur only in a very few cases, as we have already computed all the values of T and only when two successive levels are not reporting T values this contingency will arise.

10. Discussion of the results

Real time tests have shown that the software package discussed in the paper is capable of correctly decoding and checking the various types of messages and store them in the form required for machine analysis. Hydrostatic check, which has been tried on limited (Indian stations) number of stations is satisfactory.

This package however expects the data coming are in proper International format free of mutilation. This is ensured by a software package developed by Shaikh *et al.* (1975).

It is proposed to further test this package on a real time basis and to improve hydrostatic check routine. The package which is at present in Fortran IV is also proposed to be translated into 360/Assembler language to make it more efficient.

Acknowledgements

We are grateful to the Director, Northern Hemisphere Analysis Centre, New Delhi for his critical evaluation of this programme and helpful suggestions. We also thank the Director General of Observatories for his continued interest in this project. Finally, we are happy to acknowledge the help of all our colleagues and in particular to Shri Hari Kishan for preparation of the diagrams and Miss Sneh Lata Thapar for the typing work.

REFERENCES

Hinkelmann, K. H. 1969 Automatic processing of data, Lectures on numerical short range weather prediction.

Shaikh, Z. E., Dewan, B. N. and Datta, R. K. 1975 *India J. Met. Hydrol. Geophys.* **26**, 3, pp. 309-313.

APPENDIX I

SYNOP Code

ZCZC

SM or SI RUMS or DEMS etc. DTTM
MMXX YYGG

IIiii Nddff VVwwW PPPTT $N_hC_LhC_MC_H$
T_d $T_d9P_{24}P_{24}$

– – – – – – – – –

IIiii Nddff VVwwW PPPTT $N_hC_LhC_MC_H$
T_d T_d $9P_{24}P_{24}$

NNNN

PILOT Code

ZCZC

PPAA YYGGa$_4$ IIiii

$44nP_1P_1$ or $55nP_1P_1$

ddfff ddfff . . . . . ddfff
ddfff ddfff . . . . . ddfff

NNNN

TEMP Code

ZCZC

USIN DEMS DT TMTM

TTAA YYGGI$_d$ IIiii

99 $P_0P_0P_0$ $T_0T_0T_{a0}D_0D_0$ $d_0d_0f_0f_0f_0$
$P_1P_1h_1h_1h_1$ $T_1T_1T_{a1}D_1D_1$ $d_1d_1f_1f_1f_1$

– – – – – – –

$P_nP_nh_nh_nh_n$ $T_nT_nT_{an}D_nD_n$ $d_nd_nf_nf_nf_n$
88 $P_tP_tP_t$ $T_tT_tT_{at}D_tD_t$ $d_td_tf_tf_tf_t$

$\left.\begin{matrix}77\\66\end{matrix}\right\}P_mP_mP_m$ $d_md_mt_mt_mt_m$

TTAA YYGGI$_d$ IIiii

$\left.\begin{matrix}77\\66\end{matrix}\right\}P_mP_mP_m$ $d_md_mt_mt_mt_m$

NNNN

The symbols have the following meaning—

ZCZC	=	Indicator for beginning of data
SM/SI	=	Indicator for Synop data of standard/intermediate hour
DEM/RUMS/etc.	=	Designator for the meteorological service
DT	=	Data in two digits
TM	=	Time in two digits
TMTM	=	Time in four digits
MMXX	=	Synop data beginning indicator
US	=	Indicator of Temp. data
UI	=	Indicator for Pilot data
TTAA	=	Beginning of Temp. data Part I
PPAA	=	Beginning of Pilot data Part I
NNNN	=	Indicator for end of data group

Rest of the symbols have their standard meanings.

APPENDIX II

List of subroutines

RT 1	TDWDS and STOR have been called
RT 2	TDWDS have been called
RT 3	CTEND and CSTOR are called
CTEND	for calculating temperature, wind direction and speed
CSTOR	for calculating heights of different levels in China data
CHINA	Checking China data
STOR	for calculating heights of different levels
TDWDS	for calculating temperature, wind direction and speed
TDP	for calculating temperature, dew point
WDS	for calculating wind direction and speed
DTM	Checking date and time
BLINDX	Checking block and index
TEMPER	Checking Temp. data
PILOT	Checking Pilot data
CHECK	Checking duplicate stations