
INDIAN INSTITUTE OF TROPICAL METEOROLOGY

PASHAN, PUNE - 411 008

COMPUTER DIVISION

SORT MERGE UTILITY

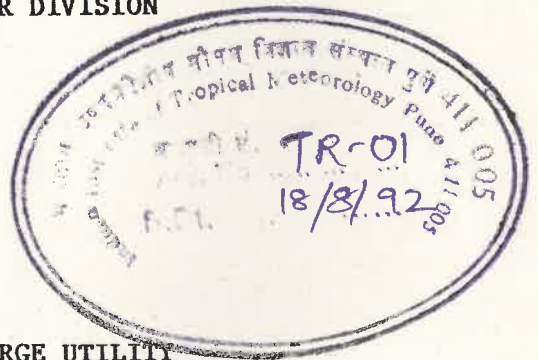
MRS. R.R. JOSHI

AUGUST 1992

INDIAN INSTITUTE OF TROPICAL METEOROLOGY

PASHAN PUNE - 411 008

COMPUTER DIVISION



SORT MERGE UTILITY

MRS. R.R. JOSHI

INDIAN INSTITUTE OF TROPICAL METEOROLOGY, PUNE 8.

01 Security classification: Unclassified 02 Distribution: Unrestricted

03 Report status: New 04 Series: CPT 05 Report No.: 01/92

06 Report Type: Technical Report-Utility

07 Title and subtitle: SORT MERGE UTILITY

08 Personal author(s): Mrs. R.R. Joshi

09 Affiliation author(s): Indian Institute of Tropical Meteorology, Pune 8.

10 Corporate author(s): ---

11 Originating unit: Computer & Data Division

12 Sponsor(s): Name: In house

13 Date of submission: August 1991 Publication/Issue Date: August 1992

14 Form of distribution: Hard copy

15 Language of text: English

16 No. of references: Three

17 Abstract: This document describes the procedures for handling SORT/MERGE utility of ND-system. The document is self explanatory and describes salient features and terms of this utility. Any user familiar with SINTRAN operating system will find it easy to use this utility by following this document.

18 Keywords: Major key, Minor key, Ascending, Descending, Record, File.

19 Whether Refereed: Yes

The information within this publication is given in good faith and considered to be true, but IITM accepts no liability for error, omission and for loss or damage arising from its use.

SORT-MERGE UTILITY

INTRODUCTION : SORT-MERGE Utility is one of the many system utilities available on ND-560/CX supermini computer system. In this report, the author has tried to present a brief description about its usage. Details are provided where ever felt necessary. An effort is also made to highlight the points which are not given in the concerned manual but are crucial in the proper use of this utility and in understanding of the relevant features. Author has tried to document this report in simple terms to facilitate its understanding by the new users.

SORT-MERGE package has ND - version E for ND-500

This utility can be used for rearranging data on mass storage files. It can be used in two ways.

1. SINTRAN III subsystem
2. Subroutine in application program.

There is no much difference in CPU time when we use sort utility by SINTRAN III system and by subroutine program. Sort means all records or lines in one input file are ordered (ascending or descending) according to the contents of one or several fields.

```

      .------.
      : Input   :
      : file   :
      -----
      !
      \!/
      .------.
      : Sort-Merge :
      : program   :
      :           :
      -----
      !
      \!/
      .------.
      : Output   :
      : file    :
      -----

```

Requirements for sort utility: Input file, Output file, File organization Record size, No. of sort fields, Position of fields in the record, Size and type of each field.

One can run sort-merge program directly from terminal by giving commands in one line or separated by carriage return.

How to use sort utility : Suppose one wants to sort a file with record length of 80 characters, text type and only on one field. Position of the field is starting from 9th column and length is 12 columns. Output file should be in ascending order. Detailed information regarding commands is given below. In case of any error repeat the command i.e preceded by * (which is a prompt for this utility) and not the parameter. HELP command gives an explanation of error messages or list of all commands and their parameters.

Help command will give the information as follows :

EXIT	
HELP	
SCRATCH-FILE	(File name)
BLOCK-FACTOR-OUT	(Number) (Records/Characters)
BLOCK-FACTOR-IN	(Number) (Records/Characters)
ALTERNATIVE-COLLATING-SEQUENCE	(File name)
RECORD-DESCRIPTION	(Rec-length) (Number of keys) (Record type)
KEY-DESCRIPTION	(Key-position) (Key length) (Sequence) (Type) (Pos) (Length)
SORT	(Input-File) (Output-File)
MERGE	(Number of Files) (Input-File),... (Output- File)
SECURE	(Off/On)
ON-SEGMENT	(Off/On)

:DATA is the default file extension.

***** INFORMATION UNDERLINED SHOULD BE GIVEN BY USER *****

@ SORT-MERGE-500

*REC-DESC

RECORD-LENGTH : 1:80

NUMBER OF KEYS : 1

RECORD TYPE : Text

*KEY-DESC

POSITION : 9

LENGTH OF FIELD : 12

SEQUENCE : Ascending

TYPE OF DATA : ASCII

*SORT

INPUT-FILE : A:symb

OUTPUT-FILE : "ASORT:Symb"

Details of Commands :

1. RECORD-DESCRIPTION : There are three parameters.

- i. Record length - It can be given as single number or two numbers separated by colon. e.g. 80 or 50:80
50 - Minimum number of characters in the record
80 - Maximum number of characters in the record
- ii. No. of keys - You can have 10 fields if sort-merge is used in interactive mode or from mode file and 99 fields in case of calling subroutine.

- iii. Record Type -
1. Text : The file is written in 7 bit ASCII character each record is terminated by CR+LF (carriage return and line feed).
 2. Fixed: All records of file are of the same length. For record length add two characters more in your record length for CR+LF. This type is five times faster than text type.

Example :

```
-----
@ SORT - MERGE - 500
*REC-DES 50:80,2,TEXT
*KEY-DES 10 30 A A
*SORT RDATA:DATA "OUT:DATA"
EXIT
```

2. KEY DESCRIPTION : There are 4 parameters. Position, length, sequence, type of data.

- i. Position : Position of the record where the field begins.
- ii. Length : Size of the field in character. This parameter can not exceed the maximum length of the record and also not more than 255 bytes.
- iii. Sequence : Ascending e.g. 1,2,3.....9
Descending e.g. 9,8,7.....1
- iv. Type of data : ASCII : field is treated according to ASCII alphabet.

Integer : field is treated as binary word,
length either 2 or 4 bytes

If there are more than one key first key from left is treated as 'Major Key' and others are treated as 'Minor Keys'. Priority is given from left to right for Minor keys. Examples can be seen on page No. 8 and 9. On page No. 8 Major key description is 18 2 A A and Minor key descriptions are 20 1 A A , 1 5 A A. On page No. 9 Major key is 1 5 A A and Minor keys are 18 2 A A , 20 1 A A.

3. SORT : There are two parameters : Input-file and output-file.

For input and output file peripheral device such as line printer, terminal can not be used. For Output-file if it is non-existing file give file name in double quotes and if it is existing file then give only file name with proper extension. After completion of successful sorting it gives the message ----records sorted. We can give same name for Input-file as well as Output-file but after sorting contents in Input-file will be replaced by Output-file (sorted file). Utility Opens and closes the files itself.

All commands can be put in one mode file and can be used to save the typing every time.

Sort utility used as subroutine called from program.

Example: There are two fields to be sorted, one field is starting from 1st column with field length 20 col. 0- ascending seq. 9- ASCII TYPE and second field starting with 25th column with 5 field length. Format for call sort is as follows:

```
CALL SORT(INPUT,OUTPUT,SCRATCH,MINLEN,MAXLEN,RECTYPE,NO. OF FIELDS,FIELD  
ARRAY,BUFFSIZE,BUFFAREA, BLOCK-IN, BLOCK-OUT, COLL-FILE, STATUS)
```

Input, Output, Scratch are some file names which are used in sort. Output file should be existing file. Before execution sort should be loaded in linkage loader. Command is LOAD SORT.

MINLEN : Minimum length in bytes of the record in the file.
MAXLEN : Maximum length in bytes of the record in the file.
RECTYPE: An integer value indicating specific type
 0- Fixed
 1- Text
NO.OF FIELDS: Maximum 99 fields can be sorted.
FIELD ARRAY : Consist of four parameters i.e. position,
 length, sequence, type
BUFFER SIZE : An integer giving the size can be declared in program.
BUFFER AREA : An integer array of the size specified in program.
BLOCK-IN : An integer specifying the size of mag. tape
 blocks in bytes used in input file.
BLOCK-OUT : An interger specifying the size of mag. tape
 blocks in bytes used in output file.
 0- tape is not used or in case of standard
 block size.
COLL-FILE : The name of a file containing the ALTERNATIVE
 - ASCII character set. If there is no
 ALTERNATIVE - ASCII field in use, this
 parameter must be zero. Default file is
 :DATA.
STATUS : An integer =0 in case of sucessful sorting
 = Non zero indicate an error.

If the value of STATUS is in the range 0-256 decimal it denotes an operating system error, usually from the file system and for detail information refer SINTRAN III REFERENCE MANUAL. If the value of STATUS from the sort subroutine is in the range of 2584 to 2604 decimal , types of errors are described on page No. 61 of ND-100/500 SORT-MERGE User Guide.

Example:

```
PROGRAM SORT1  
PARAMETER (IBUFSZ=2000B)  
CHARACTER*32 INFILE,OUTFILE  
INTEGER BUFFER(IBUFSZ)  
INTEGER IFNO,IFIELD(8)  
DATA IFNO/2/IFIELD/1,20,0,9,25,5,1,2/  
INFILE= '  
OUTFILE= '  
IRECT=1  
CALL SORT(INFILE,OUTFILE,0,80,80,IRECT,IFNO,IFIELD,IBUFSZ,BUFFER,0,0,0,  
IISTATUS)  
IF (IISTATUS.NE.0)WRITE(1,90)IISTATUS  
90 FORMAT(*ERROR IN SORT*,I6)  
END
```

Table for type of record

Code No.	Type
0	- ASCII
1	- Alternative ASCII
2	- Numeric unsigned
3	- Numeric leading separate
4	- Numeric trailing separate
5	- Numeric leading embedded
6	- Numeric trailing embedded
7	- Integer
8	- BCD
9	- ASCII upper case
10	- Bit string
11	- Real

If there is huge amount of data and we have to save the space on storage media we can use code Nos. 5 and 6. Here Numeric trailing (leading) embedded means the sign of the numeric quantity is over punched on the last (first) numeric digit. Characters corresponding to

- i) + + + + + + + + + are {,A,B,C,D,E,F,G,H,I and
 0 1 2 3 4 5 6 7 8 9
- ii) - - - - - - - - - are },J,K,L,M,N,O,P,Q,R.

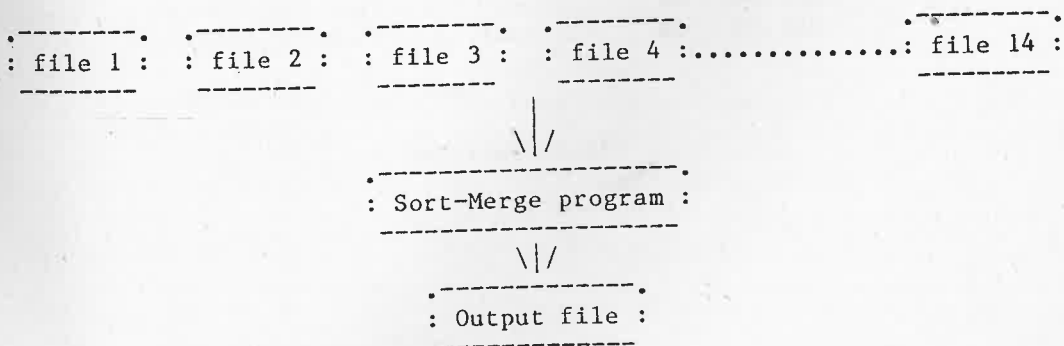
Example is as follows:

Code No.	Input Data	Data Display	Data to be punched
5	-1250	1250	J250
6	-1250	1250	125}

***** MERGE UTILITY *****

Merge : It combines two or more input files already sorted according to the contents of specified fields. The result is a single sorted file. Maximum 14 files can be merged. Each file should have same record layout, organization and sorted in the same sequence.

- Advantages : i) With one command, number of files can be merged.
 ii) When data files are extremely larger or if there is limited disk space then one can sort small separate files and then merge into one output file.



Each data set can have different frequency but it should be organized in the same way. If two input sorted files (to be merged) have some identical records then those identical records also appear in output file i.e. the multiplicity of a record occurrence is maintained.

Example:

Input file 1	Input file 2	Output file
5	5	5
		5
6		6
7	7	7
		7
8	8	8
		8

Merge : No. of input files, name of input files, output file.

No. : maximum 14

Input-file : names of the input files separated by commas or blanks

Output-file : output file must be different from all input files

Similar to sort operation after successful completion of merge it gives no. of records merged.

@ SORT-MERGE

* REC-DES : 80;1,TEXT

* KEY-DES : 11,30,A,A

* MERGE : 3 TEST1, TEST2, TEST3, FINAL-TEST

* EXIT

Acknowledgement: The author wishes to express her grateful thanks to Shri. D.R. Sikka and Shri. R. Suryanarayana, for their encouragement and providing facilities. Author wishes to thank S/Shri S.S. Aralikatti and P.W. Dixit for their valuable help. Author wishes to thank Shri M.K.Tandon for reviewing this technical report and offering constructive suggestions. The inclusion of his suggestions have led to a significant improvement in the presentation of text.

References:

ND 100/500 Sort-Merge User Guide ND - 60.236.1 EN
ND FORTRAN Reference Manual ND - 60.145
SINTRAN III Reference Manual ND - 60.128

Symbols used @ : IITM>

: Carriage return

* : Ready for commands.

40372 2913 4759881101441010710083 -99910025 9968 9913 9939 9997100631012810137
40372 2913 4759882102141017510151 -9991008910031 99761000210060101291019510206
40372 2913 4759883 118 155 197 -999 326 359 380 375 341 285 196 146
40372 2913 4759884 95 96 99 -999 93 99 105 116 92 158 103 96
40372 2913 4759885 38 8 13 -999 00 00 00 00 00 00 00 8
40427 2616 5037881101951015910134 -9991006810016 9967 999210044 -999 -99910186
40427 2616 5037882101971016110136 -9991007010018 9969 999410046 -999 -99910188
40427 2616 5037883 162 179 205 -999 301 329 339 345 330 -999 -999 196
40427 2616 5037884 131 152 159 -999 221 251 304 322 316 -999 -999 162
40427 2616 5037885 15 107 2 -999 00 00 00 00 00 -999 -999 1
40372 2913 475971110132100981009510047 9997 9944 -999 9935 9998100741011810131
40372 2913 47597211013710131100771005410030 9967 9931 994910011100731011410156
40372 2913 47597311014810107100941004810012 9954 9917 9925 9989100551012310143
40372 2913 47597411012010117100881004610011 9951 9915 9936 9993100601011310136
40372 2913 47597511014110118100921005710007 9946 -999 9926 998310058 -99910140
40372 2913 47597611013310111100931007110032 9969 9934 994810013100601011010124
40372 2913 475977110146 -999101081004110012 9951 9904 993810001100761011910134
40372 2913 47597811014910124100861005210009 9949 9915 994310000100701014010137
40372 2913 47597911014310101100861005910033 9976 9936 9942 9996100681010010129
40372 2913 47598011013510098100711004510012 9940 9915 993910004 -9991011010124
40372 2913 4759811 -99910110100741004210021 9960 9913 993510002100701010810134
40372 2913 47598211012810116100891005310031 9967 -999 993610011100671011810145
40372 2913 47598311015810123101031006210012 9977 -999 9937 9992100751012010142
40372 2913 4759841101301012310075 -99910012 9944 9923 994810002100831011510142
40372 2913 47598511012310115101011004610008 9956 9931 9938 9998100741011610139
40372 2913 47598611014410110100781004710017 9957 9933 9947100141007010112 -999
40372 2913 47598711016210117100751007310019 9974 9943 995010007100701012010147
40427 2616 50377111018510153101551010310051 9994 9971 998910053101271017310185
40427 2616 5037721101901018610139101151008710034 99821000810066101291016510179
40427 2616 5037731102071016610146101031006310002 9972 998310051101121017210195
40427 2616 5037741101781017410142101001006810005 9970 998910048101181017010188
40427 2616 50377511019310173101461011310055 9993 9973 997810035101091016710191
40427 2616 5037761101861016110141101201007310007 9978 999210055101091015710176
40427 2616 503777110191 -999101561009710063 9999 9957 999310053101241016210185
40427 2616 50377811019910179101401010610061 9997 9970 99931005010122 -999 -999
40427 2616 5037791101921015110137 -9991008410016 9989 -99910062101271016210188
40427 2616 50378011019510159101321010310070 9999 -999 -99910065101181017110185
40427 2616 5037811101831017610138101011007310021 9973 999410062101291016510193
40427 2616 5037821101881017410150 -99910086 -999 -999 -999 -999 -999 -999
40427 2616 5037831 -999 -999 -999 -999 -999 -999 -999 -999 -999 -9991018210199
40427 2616 50378411018810184101331010210067 9997 99791000110058101411016910191
40427 2616 5037851101781016910152101071006110009 9985 999610055101261016910193
40427 2616 5037861101991017210139101051006910006 99911000610073101301016510205
40427 2616 503787110218101791013710131 -99910029100001000710065101301016910199

COMMANDS FOR YEARWISE SORTING

@SORT-MERGE

--- ND-500 SORT/MERGE SYSTEM VERSION E---

*REC-DES 1:80,3,TEXT

*KEY-DES 18 2 A A,20 1 A A,1 5 A A

*SORT 88:SYMB Y88:SYMB
48 RECORDS SORTED

CPU TIME= 0.1 SECONDS

*EXIT - EXIT -

*** YEAR WISE SORTING ***

40372	2913	475971110132100981009510047	9997	9944	-999	9935	9998100741011810131
40427	2616	50377111018510153101551010310051	9994	9971	998910053101271017310185		
40372	2913	47597211013710131100771005410030	9967	9931	994910011100731011410156		
40427	2616	5037721101901018610139101151008710034	9982	1000810066101291016510179			
40372	2913	47597311014810107100941004810012	9954	9917	9925 9989100551012310143		
40427	2616	5037731102071016610146101031006310002	9972	998310051101121017210195			
40372	2913	47597411012010117100881004610011	9951	9915	9936 9993100601011310136		
40427	2616	5037741101781017410142101001006810005	9970	998910048101181017010188			
40372	2913	47597511014110118100921005710007	9946	-999	9926 998310058 -99910140		
40427	2616	50377511019310173101461011310055	9993	9973	997810035101091016710191		
40372	2913	47597611013310111100931007110032	9969	9934	994810013100601011010124		
40427	2616	5037761101861016110141101201007310007	9978	999210055101091015710176			
40372	2913	475977110146 -999101081004110012	9951	9904	993810001100761011910134		
40427	2616	503777110191 -999101561009710063	9999	9957	999310053101241016210185		
40372	2913	47597811014910124100861005210009	9949	9915	994310000100701014010137		
40427	2616	50377811019910179101401010610061	9997	9970	99931005010122 -999 -999		
40372	2913	47597911014310101100861005910033	9976	9936	9942 9996100681010010129		
40427	2616	5037791101921015110137 -9991008410016	9989	-999	10062101271016210188		
40372	2913	47598011013510098100711004510012	9940	9915	993910004 -9991011010124		
40427	2616	50378011019510159101321010310070	9999	-999	-99910065101181017110185		
40372	2913	4759811 -99910110100741004210021	9960	9913	993510002100701010810134		
40427	2616	5037811101831017610138101011007310021	9973	999410062101291016510193			
40372	2913	47598211012810116100891005310031	9967	-999	993610011100671011810145		
40427	2616	5037821101881017410150 -99910086	-999	-999	-999 -999 -999 -999 -999		
40372	2913	47598311015810123101031006210012	9977	-999	9937 9992100751012010142		
40427	2616	5037831 -999 -999 -999 -999 -999 -999	-999	-999	-999 -999 -9991018210199		
40372	2913	4759841101301012310075 -99910012	9944	9923	994810002100831011510142		
40427	2616	50378411018810184101331010210067	9997	9979	1000110058101411016910191		
40372	2913	47598511012310115101011004610008	9956	9931	9938 9998100741011610139		
40427	2616	5037851101781016910152101071006110009	9985	9996	10055101261016910193		
40372	2913	47598611014410110100781004710017	9957	9933	9947100141007010112 -999		
40427	2616	5037861101991017210139101051006910006	9991	10006	10073101301016510205		
40372	2913	47598711016210117100751007310019	9974	9943	995010007100701012010147		
40427	2616	503787110218101791013710131 -99910029	10000	10007	10065101301016910199		
40372	2913	4759881101441010710083 -99910025	9968	9913	9939 9997100631012810137		
40427	2616	5037881101951015910134 -99910068	10016	9967	999210044 -999 -99910186		
40372	2913	4759882102141017510151 -99910089	10031	9976	1000210060101291019510206		
40427	2616	5037882101971016110136 -99910070	10018	9969	999410046 -999 -99910188		
40372	2913	4759883 118 155 197 -999	326 359 380	375 341 285 196 146			
40427	2616	5037883 162 179 205 -999	301 329 339	345 330 -999 -999 196			
40372	2913	4759884 95 96 99 -999	93 99 105 116	92 158 103 96			
40427	2616	5037884 131 152 159 -999	221 251 304 322	316 -999 -999 162			
40372	2913	4759885 38 8 13 -999	00 00 00 00	00 00 00 8			
40427	2616	5037885 15 107 2 -999	00 00 00 00	00 -999 -999 1			

*** COMMANDS FOR INDEX NUMBER WISE SORTING ***

@SORT-MERGE

--- ND-500 SORT/MERGE SYSTEM VERSION E---

*REC-DES 1:80,3,TEXT

*KEY-DES 1 5 A A,18 2 A A,20 1 A A

*SORT 88:SYMB I88:SYMB

48 RECORDS SORTED

CPU TIME= 0.1 SECONDS

*EXIT

- EXIT -

*** INDEX NUMBER WISE SORTING ***

40372	2913	475971110132100981009510047	9997	9944	-999	9935	9998100741011810131
40372	2913	47597211013710131100771005410030	9967	9931	994910011100731011410156		
40372	2913	47597311014810107100941004810012	9954	9917	9925	9989100551012310143	
40372	2913	47597411012010117100881004610011	9951	9915	9936	9993100601011310136	
40372	2913	47597511014110118100921005710007	9946	-999	9926	998310058	-99910140
40372	2913	47597611013310111100931007110032	9969	9934	994810013100601011010124		
40372	2913	475977110146	-999101081004110012	9951	9904	993810001100761011910134	
40372	2913	47597811014910124100861005210009	9949	9915	994310000100701014010137		
40372	2913	47597911014310101100861005910033	9976	9936	9942	9996100681010010129	
40372	2913	47598011013510098100711004510012	9940	9915	993910004	-9991011010124	
40372	2913	4759811	-99910110100741004210021	9960	9913	993510002100701010810134	
40372	2913	47598211012810116100891005310031	9967	-999	993610011100671011810145		
40372	2913	47598311015810123101031006210012	9977	-999	9937	9992100751012010142	
40372	2913	4759841101301012310075	-99910012	9944	9923	994810002100831011510142	
40372	2913	47598511012310115101011004610008	9956	9931	9938	9998100741011610139	
40372	2913	47598611014410110100781004710017	9957	9933	9947100141007010112	-999	
40372	2913	47598711016210117100751007310019	9974	9943	995010007100701012010147		
40372	2913	4759881101441010710083	-99910025	9968	9913	9939	9997100631012810137
40372	2913	4759882102141017510151	-9991008910031	9976	1000210060101291019510206		
40372	2913	4759883	118 155 197	-999 326 359	380 375 341	285 196 146	
40372	2913	4759884	95 96 99	-999 93 99	105 116 92	158 103 96	
40372	2913	4759885	38 8 13	-999 00 00	00 00 00	00 00 8	
40427	2616	50377111018510153101551010310051	9994	9971	998910053101271017310185		
40427	2616	5037721101901018610139101151008710034	9982	1000810066101291016510179			
40427	2616	5037731102071016610146101031006310002	9972	998310051101121017210195			
40427	2616	5037741101781017410142101001006810005	9970	998910048101181017010188			
40427	2616	50377511019310173101461011310055	9993	9973	997810035101091016710191		
40427	2616	5037761101861016110141101201007310007	9978	999210055101091015710176			
40427	2616	503777110191	-999101561009710063	9999	9957	999310053101241016210185	
40427	2616	50377811019910179101401010610061	9997	9970	99931005010122	-999 -999	
40427	2616	5037791101921015110137	-9991008410016	9989	-99910062101271016210188		
40427	2616	50378011019510159101321010310070	9999	-999	-99910065101181017110185		
40427	2616	5037811101831017610138101011007310021	9973	999410062101291016510193			
40427	2616	5037821101881017410150	-99910086	-999	-999 -999 -999	-999 -999	
40427	2616	5037831	-999 -999 -999 -999 -999 -999	-999	-999 -999 -999	1018210199	
40427	2616	50378411018810184101331010210067	9997	99791000110058101411016910191			
40427	2616	5037851101781016910152101071006110009	9985	999610055101261016910193			
40427	2616	5037861101991017210139101051006910006	99911000610073101301016510205				
40427	2616	503787110218101791013710131	-99910029100001000710065101301016910199				
40427	2616	5037881101951015910134	-9991006810016	9967	999210044	-999 -99910186	
40427	2616	5037882101971016110136	-9991007010018	9969	999410046	-999 -99910188	
40427	2616	5037883	162 179 205	-999 301 329	339 345 330	-999 -999 196	
40427	2616	5037884	131 152 159	-999 221 251	304 322 316	-999 -999 162	
40427	2616	5037885	15 107 2	-999 00 00	00 00 00	-999 -999 1	



महाराष्ट्र उष्ण प्रशैलीय मौसम विज्ञान संस्थान पुणे 411 005
Tropical Institute of Tropical Meteorology Pune 411005

आर. नं. TR-01
दि. 18/8/92