

Migration from TAC to BUFR for TEMP observations

Mohamed Anis Satouri, December 2017
National Institute of Meteorology, Tunisia
anissatouri@gmail.com

In collaboration with Alena Trojakova
Czech Hydrometeorological Institute

Table of Contents

INTRODUCTION	2
I. Radiosonde data.....	2
II. Processing and results	3
CONCLUSION	7
APPENDIX	9

INTRODUCTION

This report summarizes the activities done during a stay at CHMI on the Migration from traditional alphanumeric codes TAC to BUFR for radiosonde observations, the priority is to progress with validation of BATOR handling of TEMP data in BUFR format as already operational in Meteo France.

The Migration to Table-Driven Code Forms (TDCF) was officially launched by WMO in 2002. November 2014 was set as the deadline of migration, but there are some issues to be solved to reach completion of the migration, where all centers can operate based on the TDCF information flow. The main issues for upper-air observations are the availability of observation reports in TDCF and the quality of data content, especially for reports separated into parts and errors associated with reformat from TAC to TDCF.

I. Radiosonde data

TAC TEMP coding

The radiosonde reporting in TAC (TEMP) was designed to fulfill requirements in the mid-20th century, such as the need to keep telecommunication messages as short as possible, which leads to some limits like the use of ‘significant’ levels to summarize the entire profile and the reporting of temperature to a precision of 0.2°.

The TAC TEMP code contains four parts (TTAA, TTBB, TTCC, TTDD): part A (standard levels: 1000, 925, 850, 700, 500, 400, 300, 250, 200, 150, and 100 hPa, plus surface, tropopause, and maximum wind levels), part B (significant levels up to 100 hPa) these ‘significant’ levels correspond principally to turning points, part C (standard levels: 70, 50, 30, 20, and 10 hPa), and part D (significant levels above 100 hPa). Complete ascents (all parts) typically have 60–120 levels, but there can be more than 200.

BUFR TEMP CODING

BUFR is only one of two Table Driven Code Formats – the second one is CREX which is a character version but in practice there seems to be little or no use of CREX.

The table driven codes BUFR (Binary Universal Form for the Representation of meteorological data) offer great advantages in comparison with the traditional alphanumeric codes. The main features of these codes are self-description, flexibility and expandability, which are fundamental in times of fast scientific and technical evolution.

BUFR does not have parts, yet it does allow for the dissemination of much higher vertical resolution with the reporting of the time and position at each level.

Because of the complexity and the number of different data producers, the migration is taking longer than anticipated; according to the WMO’s timetable, TEMP reports should have ceased in November 2014. In practice, only a small proportion of alphanumeric reports have ceased so far, but this proportion will increase over time.

Two types of radiosonde reports in BUFR format are being exchanged globally:

- BUFR reports generated directly from the radiosonde data and known as native BUFR reports. These native BUFR reports are often high resolution.
- BUFR reports generated through the conversion of TEMP reports and known as reformatted TEMP. These reformatted TEMP are mostly still in four parts (corresponding to TEMP parts A–D); such reports do not meet the BUFR coding regulations and are difficult for NWP centers to use. Reformatted reports cannot take advantage of the higher precision and other advantages of BUFR.

Balloon Drift

According to [1], the drift of radiosonde balloons during their ascent has generally been considered a negligible factor in applications involving radiosonde observations. Typical drift distances are a few kilometers in the lower troposphere, ~5 km in the midtroposphere, ~20 km in the upper troposphere, and ~50 km in the lower stratosphere, although there is considerable variability due to variability in climatological winds. Drift distances tend to increase with height above the surface, be larger in midlatitudes than in the tropics, be larger in winter than in summer, and vary with wind (and consequent balloon drift) direction. Most estimates of elapsed time from balloon launch to various pressure levels, due to vertical balloon rise, have median values ranging from about 5 min at 850 hPa to about 1.7 h at 10 hPa, with ranges of about 20% of median values. Observed elapsed times exceed those estimated using assumed 5 or 6 m/s rise rates.

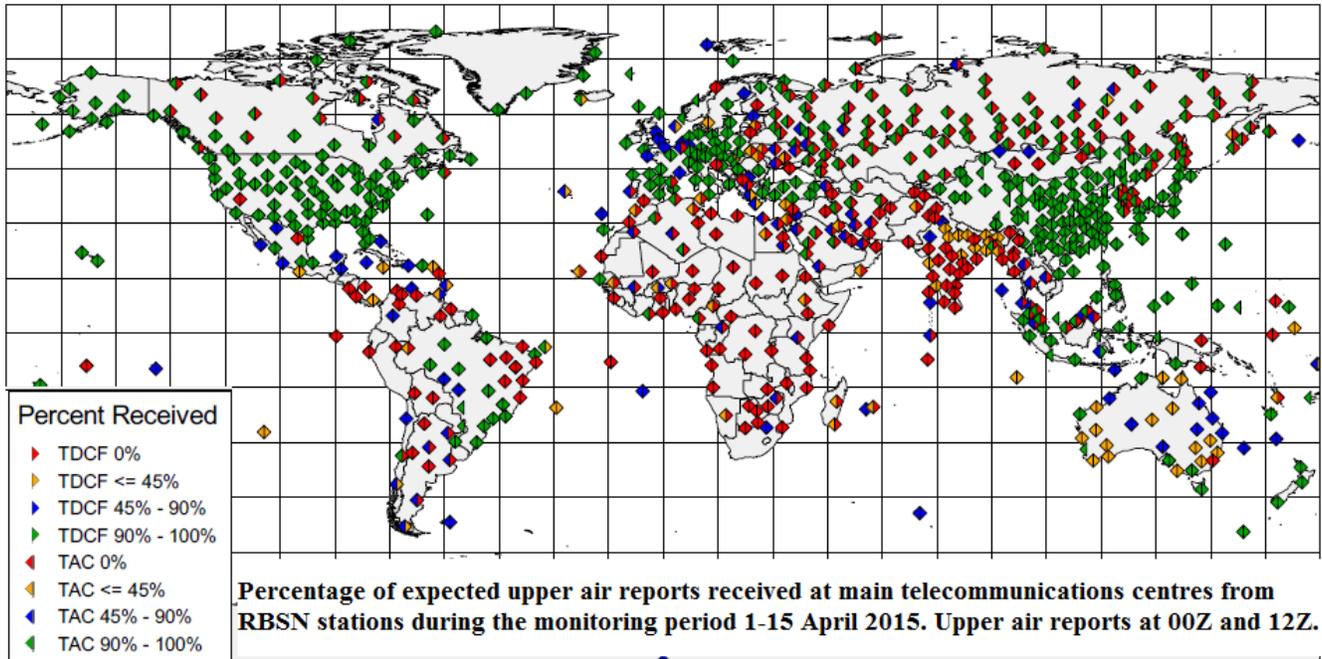


Fig.1 : Radiosonde data coverage [10]

II. Processing and results

Most of the ALADIN members' local data assimilation systems still rely on a conversion from the ASCII OBSOUL data format, the idea is to take advantage of the developments recently done inside BATOR which is now able to handle data in BUFR format.

BATOR is a software developed at Météo France to generate the ODB (Observational DataBase) for the ARPEGE/ALADIN/AROME analysis system, beside data conversion to ODB format the BATOR program can perform blacklisting, geographical selection, setting up of observation errors, etc. It was back-phased from CY43T1 to CY38T1.

Practically the following routines were back-phased (located in odb/pandor/module):

- bator_datetime_mod.F90
- bator_decodbufr_mod.F90
- bator_ecriptions_mod.F90
- bator_init_mod.F90
- bator_lectures_mod.F90
- bator_module.F90
- bator_saisies_mod.F90

For validation, messages contents were verified before and after BATOR. The following tools have been used:

ecCodes version 0.13 (bufr_ls, bufr_dump, ...)

odbsql for example:

```
odbsql -q 'select statid,date,time,lat,lon,varno,vertco_reference_1,obsvalue from hdr,body sort by varno' > log
```

Three new templates (reported in APPENDIX) were identified and introduced in param.cfg which is a configuration file to decode bufr-files.

Basically to introduce a new template, a complete user guide by Frank Guillaume could be found under this link: http://www.umn-cnrm.fr/gmapdoc/IMG/pdf/param_bator_cy43.pdf

Another way to work around is a kind of reverse engineering, it consists on:

- Decode the BUFR TEMP observation which present a new template (causing BATOR crush).
- Compare the new template to an existing one (focus on codage fields).
- Introduce the new codage fields and the new template.

The back-phased BATOR was able to read the BUFR messages and to process them properly; Also ODB produced through BUFR and through OBSOUL were compared.

The new version of BATOR introduce 2 major functionalities for the handling of radiosonde observations:

- The possibility to choose one of 3 profiles
- Thinning of observations

The 3 profiles are controlled through the namelist via the following keys:

➤ TEMPSONDSPLIT : Boolean: Used to split a radiosonde to get a vertical profile by timeslot if its value is .TRUE. Default value: .FALSE.

If the variable TempSondOrTraj = .FALSE., Has no effect

➤ TempSondOrTraj : Boolean: Allows to keep a radiosonde as a vertical profile (.TRUE.) Or to split the message into as many observations as there are levels (.FALSE). Default value: .TRUE.

The namelist key which control the thinning is NFREQVERT_TPHR. Default value: 400

Simple profile: single vertical profile	TEMPSONDSPLIT= .FALSE. TempSondOrTraj= .TRUE.
Profile with splitting	TEMPSONDSPLIT= .TRUE. TempSondOrTraj= .TRUE.
Trajectory profile	TEMPSONDSPLIT= .FALSE. TempSondOrTraj= .FALSE.

- ♦ The simple profile is a single vertical profile; all observations are reported at the same lat,lon coordinates (the position of the station).
- ♦ The profile with splitting is composed of several small vertical profiles.
- ♦ The trajectory profile presents exact time and position at each level.

After BATOR processing, the different obtained ODBs (ECMA) from the three profiles present:

- Same number of levels.
- Higher vertical resolution than (ECMA) obtained from OBSOUL.
- But less than the original bufr TEMP. (because of thinning)

Radiosonde Observations (ECMA)

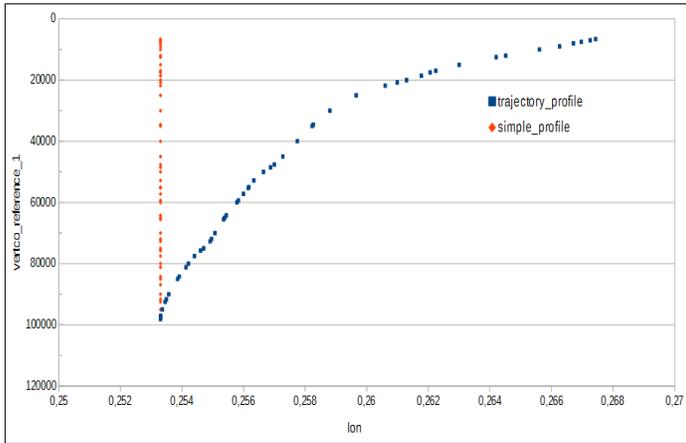


Fig.2 : Radiosonde Observations (ECMA) : simple profile VS trajectory profile

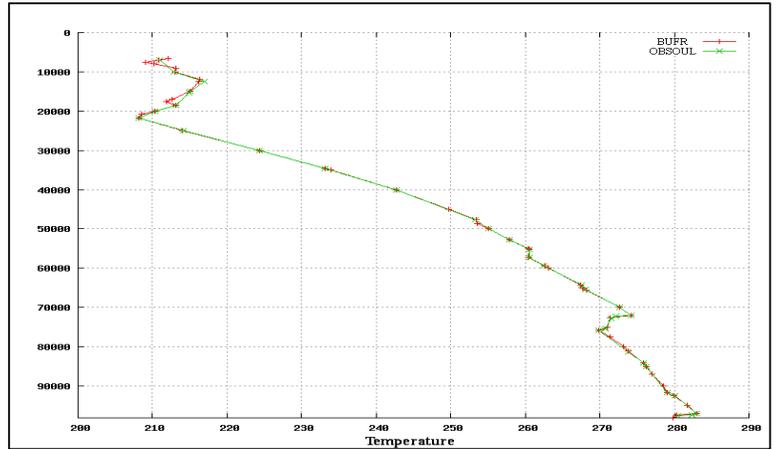


Fig.3 : Temperature profile of Radiosonde Observations (ECMA) obtained from OBSOUL (green) VS BUFR (red)

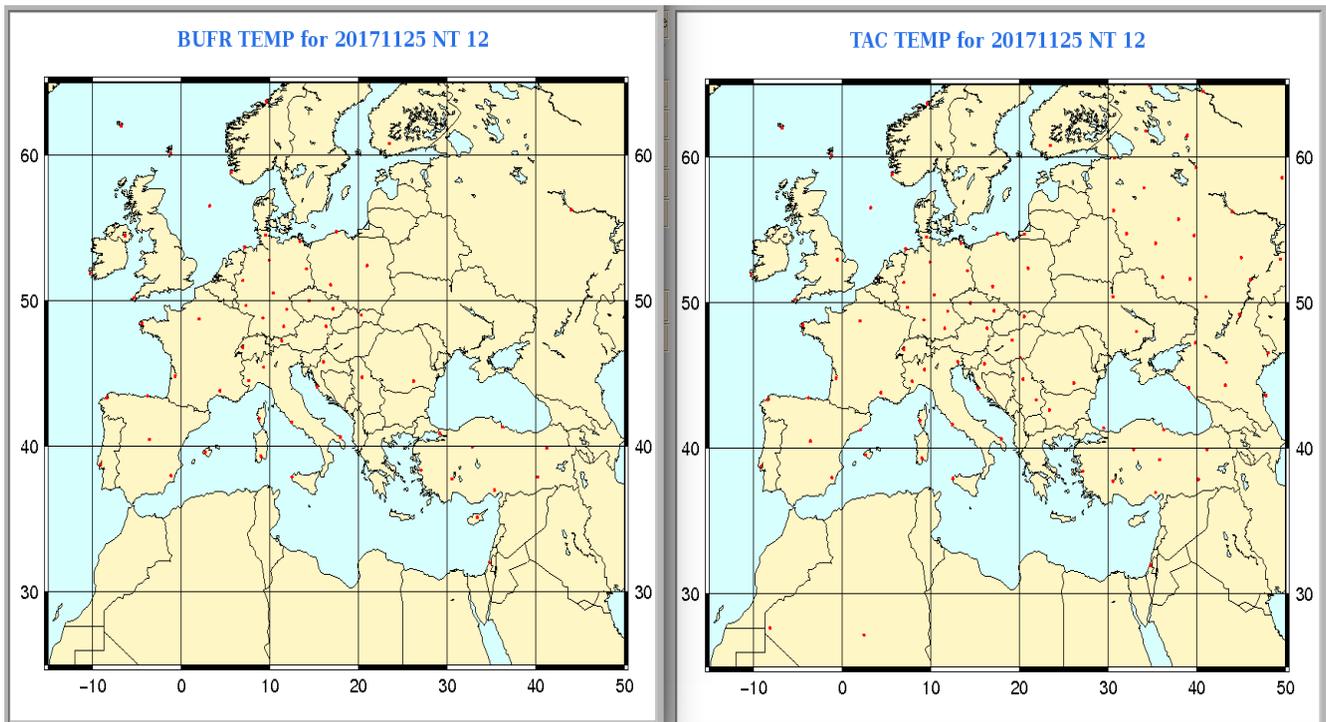


Fig.4: Available BUFR TEMP vs TAC TEMP for 20171125

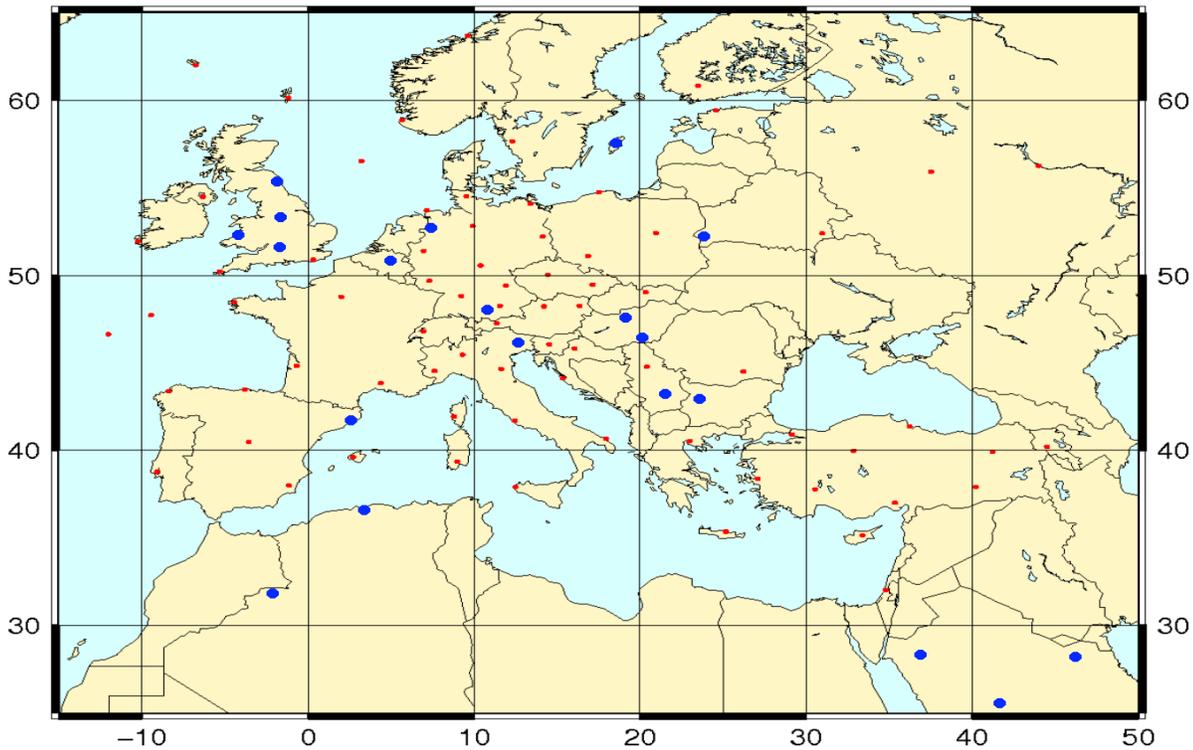


Fig.5: Missing BUFR TEMP observations (blue) retrieved from GTS in comparison with available BUFR TEMP as reported by ECMWF for 20171125

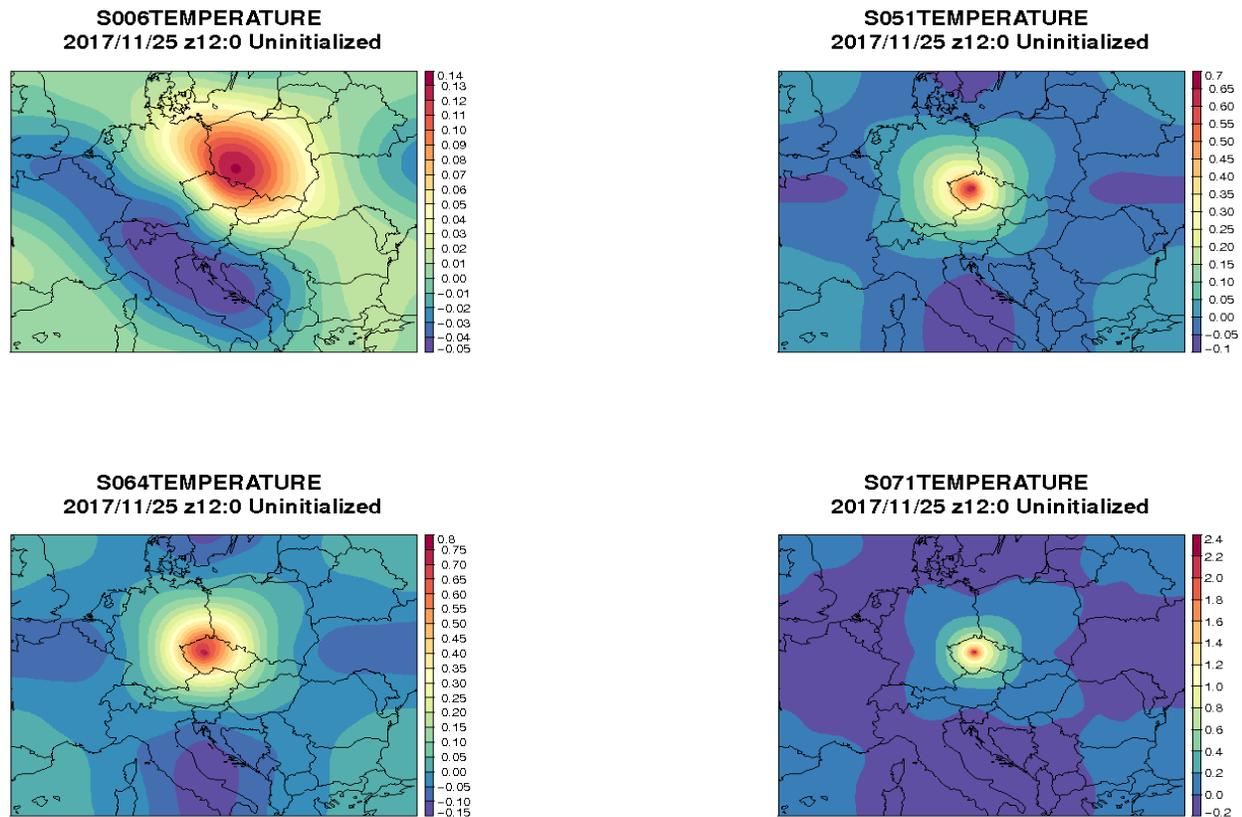


Fig.6: Temperature observation increments (observation minus background) when Radiosonde is assumed to be a single vertical profile.

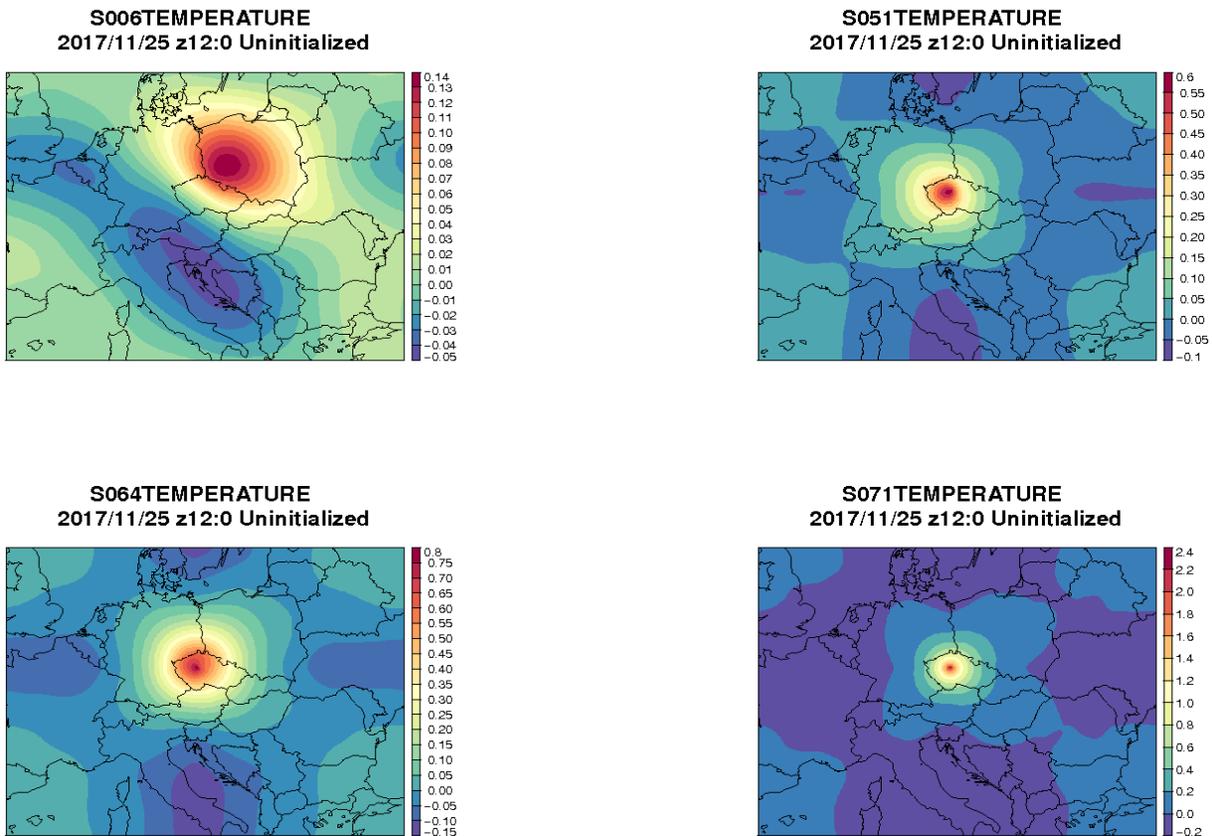


Fig.7: Temperature observation increments (observation minus background) when Radiosonde is assumed to be a trajectory profile.

First results are underwhelming; the above 2 figures didn't show big differences, despite that the large-scale mixing is switched off, observation increments when radiosonde is assumed to be a single vertical profile or trajectory profile are similar, we were expecting bigger differences, this could be explained because of the crude thinning of data (we didn't use all the available information from the observation; in a tested bufr TEMP observation which contains 1679 levels only 242 levels were kept after BATOR processing), anyway, cleaner comparisons and more testing of thinning algorithm are needed to determine the optimal way to assimilate Radiosonde observations.

CONCLUSION

Despite their relative sparsity radiosonde reports are still important for Numerical Weather Prediction (NWP) because they provide complete vertical profiles through much of the atmosphere and act as reference data for aircraft and satellite reports, but there is still much work to do to ensure a successful migration and to make good use of the extra information available in BUFR.

Hopefully, the impact of radiosonde data is expected to increase further once the reporting of the complete time and position information as well as high vertical resolution have been widely adopted and used in assimilation.

Furthermore, to control the quality of the forecasts, the software verification tool should be upgraded to be able to handle High-Resolution, Real-Time Radiosonde data.

Acknowledgment

At the end of this report I want to thank Alena Trojakova for her valuable guidance, advices and entire support during my stay, as well many thanks to the entire NWP team in Prague for their warm welcome and hospitality.

Here are some links that may be of interest:

- [1] : Global radiosonde balloon drift statistics
<http://onlinelibrary.wiley.com/doi/10.1029/2010JD014891/full>
- [2] : Progress toward High-Resolution, Real-Time Radiosonde Reports
<http://journals.ametsoc.org/doi/pdf/10.1175/BAMS-D-15-00169.1>
- [3] : Maria Monteiro: Upgrade of the source code BATOR to WMO AMDAR template 311010v7
http://www.rclace.eu/File/Data_Assimilation/2017/MMonteiro_short_stay_2017_v20170131_cor.pdf
- [4] : Maria Monteiro, 2016: Validation of a back-phased version of source code BATOR
http://www.rclace.eu/File/Data_Assimilation/2016/MMonteiro_short_stay_2016_v20160909.pdf
- [5] : VOLUME C1 - CATALOGUE OF METEOROLOGICAL BULLETINS
http://www.wmo.int/pages/prog/www/ois/Operational_Information/VolC1_en.html
- [6] : odb varno obsvalue
<http://apps.ecmwf.int/odbgov/varno/>
- [7] : Regulations for reporting TEMP, TEMP SHIP and TEMP MOBIL data in TDCF
http://www.wmo.int/pages/prog/www/WMOCodes/WMO306_vI2/LatestVERSION/BC25-TEMP_en.pdf
- [8] : keys for bufr_filter rules_file
https://software.ecmwf.int/wiki/display/ECC/WMO%3D13+element+table#WMO=13elementtable-CL_7
- [9] : Current BUFR availability and quality
<https://software.ecmwf.int/wiki/display/TCBUF/>
- [10] : Percentage of expected upper air reports received at main telecommunications centers
ftp://ftp.wmo.int/GTS_monitoring/SMM/From_WMO/sm15401.015/ANALYSIS/TEMP/tcdf_vs_tac_upper_201504.pdf

APPENDIX

A sample of BATOR for handling BUFR observations (it contains scripts, namelists and reference data) is available on beaufix under this path: /home/gmap/mrpe/trojakova/cy40t1/sample_buf/

A full description of BATOR script for cy40t1 can be found on LACE forum:
<http://www.rlace.eu/forum/viewtopic.php?f=37&t=117&start=30#p1881>

Following the 3 added templates to param.cfg

```
# template 1
BEGIN temp
14 0 1 21
codage 1 309052
codage 2 -001081
codage 3 -001082
codage 4 -002017
codage 5 -002067
codage 6 -002080
codage 7 -002081
codage 8 -002082
codage 9 -002084
codage 10 -002095
codage 11 -002096
codage 12 -002097
codage 13 -002191
codage 14 -025061
offset 1 10
values 1 001001 WMO block number
values 2 001002 WMO station number
values 3 001011 station identifier
values 4 002011 Radiosonde type
values 9 004001 Year
values 15 005001 Latitude (high accuracy)
values 16 006001 Longitude (high accuracy)
values 17 007030 Height of station ground above mean sea level
values 19 007007 Height
values 29 031002 Extended delayed descriptor replication factor
values 28 022043 Sea/water temperature
values 30 004086 Long time period or displacement
values 31 008042 Extended vertical sounding significance
values 32 007004 Pressure
values 33 010009 Geopotential height
values 34 005015 Latitude displacement (high accuracy)
values 35 006015 Longitude displacement (high accuracy)
values 36 012101 Temperature/dry-bulb temperature
values 37 012103 Dew-point temperature
values 38 011001 Wind direction
values 39 011002 Wind speed
END temp
```

```
# template 2
BEGIN temp
10 0 1 21
codage 1 309052
codage 2 -001081
codage 3 -001082
codage 4 -002017
codage 5 -002067
codage 6 -002095
codage 7 -002096
codage 8 -002097
codage 9 -002192
codage 10 -025061
offset 1 10
values 1 001001 WMO block number
values 2 001002 WMO station number
values 3 001011 station identifier
values 4 002011 Radiosonde type
values 9 004001 Year
values 15 005001 Latitude (high accuracy)
values 16 006001 Longitude (high accuracy)
values 17 007030 Height of station ground above mean sea level
values 19 007007 Height
values 29 031002 Extended delayed descriptor replication factor
values 28 022043 Sea/water temperature
values 30 004086 Long time period or displacement
values 31 008042 Extended vertical sounding significance
values 32 007004 Pressure
values 33 010009 Geopotential height
values 34 005015 Latitude displacement (high accuracy)
values 35 006015 Longitude displacement (high accuracy)
values 36 012101 Temperature/dry-bulb temperature
values 37 012103 Dew-point temperature
values 38 011001 Wind direction
values 39 011002 Wind speed
END temp
```

```
# template 3
BEGIN temp
15 0 1 21
codage 1 309052
codage 2 -001081
codage 3 -001082
codage 4 -002067
codage 5 -002080
codage 6 -002081
codage 7 -002082
codage 8 -002084
codage 9 -002095
codage 10 -002096
```

```
codage 11 -002097
codage 12 -002017
codage 13 -002191
codage 14 -025061
codage 15 -205060
offset 1 10
values 1 001001 WMO block number
values 2 001002 WMO station number
values 3 001011 station identifier
values 4 002011 Radiosonde type
values 9 004001 Year
values 15 005001 Latitude (high accuracy)
values 16 006001 Longitude (high accuracy)
values 17 007030 Height of station ground above mean sea level
values 19 007007 Height
values 29 031002 Extended delayed descriptor replication factor
values 28 022043 Sea/water temperature
values 30 004086 Long time period or displacement
values 31 008042 Extended vertical sounding significance
values 32 007004 Pressure
values 33 010009 Geopotential height
values 34 005015 Latitude displacement (high accuracy)
values 35 006015 Longitude displacement (high accuracy)
values 36 012101 Temperature/dry-bulb temperature
values 37 012103 Dew-point temperature
values 38 011001 Wind direction
values 39 011002 Wind speed
END temp
```