

GRAS SAF Report 08
Ref: SAF/GRAS/METO/REP/GSR/008
Web: www.grassaf.org
Date: 22 June 2009

The EUMETSAT
Network of
Satellite Application
Facilities



GRAS SAF Report 08

ROPP thinner algorithm

Huw Lewis

Met Office, UK

Document Author Table

	Name	Function	Date	Comments
Prepared by:	H. Lewis	GRAS SAF Project Team	22 June 2009	
Reviewed by:	D. Offiler	GRAS SAF Project Team	22 June 2009	
Approved by:	K.B. Lauritsen	GRAS SAF Project Manager	22 June 2009	

Document Change Record

Issue/Revision	Date	By	Description
Version 1.0	08 June 2009	HL	Preliminary release
Version 1.1	22 June 2009	HL	Reviewed release

GRAS SAF Project

The GRAS SAF is a EUMETSAT-funded project responsible for operational processing of GRAS radio occultation data from the Metop satellites. The GRAS SAF delivers bending angle, refractivity, temperature, pressure, and humidity profiles in near-real time and offline for NWP and climate users. The offline profiles are further processed into climate products consisting of gridded monthly zonal means of bending angle, refractivity, temperature, humidity, and geopotential heights together with error descriptions.

The GRAS SAF also maintains the Radio Occultation Processing Package (ROPP) which contains software modules that will aid users wishing to process, quality-control and assimilate radio occultation data from any radio occultation mission into NWP and other models.

The GRAS SAF Leading Entity is the Danish Meteorological Institute (DMI), with Cooperating Entities: i) European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, United Kingdom, ii) Institut D'Estudis Espacials de Catalunya (IEEC) in Barcelona, Spain, and iii) Met Office in Exeter, United Kingdom. To get access to our products or to read more about the project please go to <http://www.grassaf.org>.

Contents

1	Background	1
2	ROPP Thinner Implementation	2
2.1	Thinning method outline (<code>ropp_io_thin</code>)	4
2.2	Thinner algorithm selection <code>ropp_io_thin_select</code>	6
2.3	Thinning by sub-sampling <code>ropp_io_thin_skip</code>	6
2.4	Thinning by interpolation <code>ropp_io_thin_fixed</code>	7
2.4.1	Pre-smoothing	8
2.4.2	Interpolation	8
2.5	Limitations	9
	Bibliography	11
A	Thinning Levels	12

1 Background

The `ropp_io` module within ROPP contains routines to enable thinning of Level 1b (bending angle), Level 2a (refractivity) and Level 2b (meteorological variables) data. Thinning aims to reduce the amount of data without reducing the information content. GRAS SAF Report 01 presented some possible methods for thinning 1-dimensional (profile) data (1). It was discussed that a typical bending angle profile does not contain more than about 250 pieces of information, and it is therefore desirable to reduce the data volume to about this number from about 5000 measurements obtained during an occultation. A list of 247 set impact heights were defined for use in the GRAS SAF thinned BUFR products (see Appendix A) to meet a strong user requirement for NRT data on set levels. These heights are based on the criteria of having a maximum of four points per Fresnel diameter. It is possible to perform thinning using the ROPP software for any pre-defined set of impact heights.

It was required that the ROPP thinner must:

- reduce data volume with minimal loss of information content
- not introduce any artifact such as vertical correlations
- allow BUFR encoding such that message lengths are compatible with dissemination over the GTS, In practice this means messages less than 15,000 bytes
- run quickly and add no significant overhead to the encoding itself — i.e. process a profile in seconds (or less) rather than minutes
- output on fixed levels, preferably those defined in Appendix A
- be flexible enough that different sets of arbitrary levels can be user-defined
- implement alternative, user-selectable methods

The ROPP thinner is currently implemented within the EUMETSAT Level 1¹ and GRAS SAF NRT Level 2 processing chain using the LOG thinning method and the set of impact height levels defined in Appendix A.

¹The EUMETSAT implementation is based on the same algorithms as the ROPP code, but is not identical; however the produce the same output values.

2 ROPP Thinner Implementation

ROPP supports several different thinning algorithms.

- NONE : no thinning (default).
- SAMPLE : Simple sub-sampling (select or reject 1-in-N) to no more than the maximum number of thinned levels.
- LIN : Linear interpolation to fixed number of thinned levels (no smoothing).
- LOG : Logarithmic interpolation to fixed number of thinned levels (no smoothing).
- SGLIN : Savitzky-Golay smoothing filter with linear interpolation to fixed thinned levels.
- SGLOG : Savitzky-Golay smoothing filter with log interpolation to fixed thinned levels.
- ASGLIN : Adaptive S-G smoothing filter with linear interpolation to fixed thinned levels.
- ASGLOG : Adaptive S-G smoothing filter with log interpolation to fixed thinned levels.

The thinner may be called from user programs as

```
USE ropp_io
TYPE(ROprof) :: ro_data
...
CALL ropp_io_thin(ro_data, ThinFile, DEBUG)
```

where `ThinFile` is the name of a thinning control file. This is a plain-text file which should be formatted as follows.

```
Title=<title>           ! free-text description
Method=<method>        ! Thinning method selected from options above
Nlevels=<nlev>         ! Maximum number of levels for output (between 1-5000)
Hlevels=              ! Required set of fixed levels as impact heights
  <level1>
  <level2>
  ...
  <levelnlev>
```

A number of sample thinning control files are provided within the `ropp_io/data/` sub-directory in the ROPP distribution. The impact heights on which data are to be thinned are not used by the `SAMPLE` method. Values should be unique and monotonically increasing.

The `DEBUG` argument is a logical to specify whether diagnostics on the thinning process are printed. If `TRUE`, diagnostics on the thinning process are printed to `stdout`.

The ROPP thinner may be optionally called by the `ropp2ropp`, `ucar2ropp` and `ropp2bufr` tools. This is achieved by specifying the `ThinFile` with the `'-p'` command line argument. By default in `ropp2bufr`, the `SAMPLE` method is applied to output no more than 375 points per profile. By default in `ropp2ropp` and `ucar2ropp`, no thinning is applied. Further details are available in the relevant tool man pages.

```
> ropp2ropp input_data.nc -o thinned_data.nc -p ThinFile.dat
> ucar2ropp input_ucar_data.nc -o thinned_data.nc -p ThinFile.dat
> ropp2bufr input_data.nc -o thinned_data.bufr -p ThinFile.dat
```

The ROPP thinner is currently implemented within the GRAS SAF NRT processing chain using the `LOG` thinning method and the fixed set of impact height levels defined in Appendix A.

The following variables (together with their error and quality flag values) held in the `ROpprof` ROPP data structure (2) are thinned.

Level 1b	L1 bending angle as a function of impact parameter	<code>bangle_L1</code>
	L1 impact parameter	<code>impact_L1</code>
	L2 bending angle as a function of impact parameter	<code>bangle_L2</code>
	L2 impact parameter	<code>impact_L2</code>
	Corrected bending angle as function of impact parameter	<code>bangle</code>
	Impact parameter for corrected bending angle	<code>impact</code>
	Optimised bending angle as function of impact parameter	<code>bangle_Opt</code>
	Impact parameter for optimised bending angle	<code>impact_opt</code>
	Tangent point longitude as a function of impact parameter	<code>lon_tp</code>
	Tangent point latitude as a function of impact parameter	<code>lat_tp</code>
	Tangent point azimuth as a function of impact parameter	<code>azimuth_tp</code>
Level 2a	Refractivity as a function of geometric height (computed from fixed impact levels)	<code>refrac</code>
	Observation geopotential heights	<code>geop_refrac</code>
	Observation geometric heights	<code>alt_refrac</code>
Level 2b	Temperature as a function of geopotential height	<code>temp</code>
	Specific humidity as a function of geopotential height	<code>shum</code>
	Pressure as a function of geopotential height	<code>pres</code>
	N.B. Level 2b data only thinned if Level 2a data also thinned	

Note that the ROPP thinner does not thin Level 1a data (amplitude and excess phase as a function of time). If thinning methods LOG, SGLOG or ASGLOG are selected, only those profile parameters that behave logarithmically with height (bending angle, refractivity, humidity) are interpolated logarithmically; all other parameters are interpolated in linear mode.

If there are fewer than `nlev` observation levels in the input data, there are insufficient data to thin. To guarantee the required number of output levels are still output however, an output structure filled with missing data values of size `nlev` is created. The height variables are set equal to the fixed thinned level heights as required.

2.1 Thinning method outline (`ropp_io_thin`)

1. Check that data are within valid range limits, by calling `ropp_io_rangecheck`,
2. Read and parse thinning control file
 - Determine required thinning method
 - Read number of levels (maximum or fixed, depending on the method)
 - Read the impact height values of the fixed levels if appropriate,
3. Write thinner method and thinner software version to `ROprof` structure for later output to file,
4. If thinning method is NONE or the thinning control file cannot be read, exit thinner routines,
5. Thin bending angles as a function of impact parameter.
 - Compute impact parameter from fixed impact parameter heights by adding radius of curvature R_c .

$$\text{ThinLev} = \text{ThinLev} + \text{ROdata}\% \text{GEOref}\% \text{RoC}$$
 - Ensure there are sufficient observed data points to require thinning. If fewer observation levels than thinned levels, set output to missing.
 - Thin all variables in the `ROdata%Levlb` sub-structure by the required method. This is achieved by calling `ropp_io_thin_select`.
 - For thinning tangent point longitude, if the difference between any adjacent longitude values exceeds 300 degrees (i.e. if interpolating across the ± 180 degree longitude), first transform all data from range ± 180 to 0-360 degrees. Then thin data and finally transform thinned longitude results back to ± 180 degree range.

6. Thin refractivity as a function of geometric height with respect to geoid,

- Ensure there are sufficient observed data points to require thinning. If fewer observation levels than thinned levels, set output to missing.
- If thinning to a fixed number of output impact parameter levels, convert observation geometric heights (with respect to geoid) to impact parameter.

Ensure valid refractivity values at all observation heights i , scanning profile from top to bottom.

- If valid observed refractivity at level i , $N = \text{ROdata}\%Lev2a\%refrac(i)$
- If invalid refractivity at level i , interpolate from previous good value,

$$N = \exp [\ln(N_{\text{good}}) + (H_{\text{good}} - H)(\ln(300) - \ln(N_{\text{good}}))/H_{\text{good}}]$$

where N_{good} is the previous good value at valid data level H_{good} and H is the assumed geometric height at level i , inferred from $\text{impact}(i)$. A surface refractivity value of 300 is assumed. At the highest level, N_{good} is initialised as 10^{-6} .

- Note that interpolation using assumed values is only performed to give sensible impact parameter values for all fixed levels, which are later inverted back to geometric heights. By definition, the refractivity at these heights is missing, so these levels won't be useful anyway.

$\text{Lev}(i) =$

$$(1+10^{-6}N(i)) (\text{Lev2a}\%alt_refrac(i) + \text{GEOfref}\%RoC + \text{GEOfref}\%Undulation)$$

- Thin all variables in the $\text{ROdata}\%Lev1b$ sub-structure by the required method. This is achieved by calling `ropp_io_thin_select`.
- If thinned to a fixed number of output impact parameter levels, convert thinned impact parameter levels back to geometric height (with respect to geoid).

Ensure valid refractivity values at all thinned heights i , scanning profile from top to bottom.

- If valid thinned refractivity at level i , $N = \text{ROdata}\%Lev2a\%refrac(i)$
- If invalid refractivity at level i , interpolate from previous good value,

$$N = \exp [\ln(N_{\text{good}}) + (H_{\text{good}} - H)(\ln(300) - \ln(N_{\text{good}}))/H_{\text{good}}]$$

where N_{good} is the previous good value at valid data level H_{good} and H is the thinned impact height at level i , inferred from $\text{impact}(i)$. A surface

refractivity value of 300 is assumed. At the highest level, N_{good} is initialised as 10^{-6} .

- Note interpolation is only performed to provide realistic height values where thinned refractivity is missing, the returned refractivity values will still be missing.

$$\text{ThinLev}(i) = (\text{ThinLev}(i)/(1+10^{-6}N(i))) - \text{GEOref}\%RoC - \text{GEOref}\%Undulation$$

7. Thin meteorological variables (temperature, humidity, pressure) as a function of geopotential height with respect to geoid. Note this step is only possible if refractivity data were thinned. Otherwise, there is no way to convert the `ThinLev` impact parameters to geometric heights and hence to geopotential heights, so the data cannot be thinned.
 - Ensure there are sufficient observed data points to require thinning. If fewer observation levels than thinned levels, set output to missing.
 - Compute thinned geopotential height levels from thinned geometric height (calculated in thinning refractivity data) using `ropp_utils` function `geometric2geopotential`.
 - Set `ROdata%Lev2b%geop` equal to the thinned geopotential height levels.
 - This all other variables in the `ROdata%Lev2b` sub-structure by the required method. This is achieved by calling `ropp_io_thin_select`.
8. Check that all thinned data are within valid range limits, by calling `ropp_io_rangecheck`.

2.2 Thinner algorithm selection `ropp_io_thin_select`

The thinning method required as set by the input `Method` string defined in the thinning control file is called from the subroutine `ropp_io_thin_select`. Only those input data where both the observation level and the observation data are valid are input to the appropriate thinning algorithm. At least two good values on valid levels are required, else all thinned values are returned as missing on the required thinned levels.

2.3 Thinning by sub-sampling `ropp_io_thin_skip`

Thinning by sub-sampling (`SAMPLE`) is implemented by calling routine `ropp_io_thin_skip`, and then selecting every `skip1th` sample from the input observed profile or rejecting every

skip2th sample of those selected as required. Parameters skip1 and skip2 are computed for a given input profile by ropp_io_thin_skip.

```
call ropp_io_thin_skip ( nLev, nThinLev, skip1, skip2, nSamp, DEBUG)

do i = 1, nLev, skip1 ! select every skip1-th sample
  if ( mod( i, skip2) == 0 ) cycle      ! reject every skip2-th sample
  ....
enddo
```

Routine ropp_io_thin_skip caters for thinning in cases where the input profile is high resolution (e.g. raw sampling) or low resolution (e.g. pre-thinned) but with too many levels for output to BUFR for example.

If the number of input levels nLev is more than twice the maximum requested number of output thinned levels nThinLev, the algorithm selects 1 in N levels (giving a total of nSamp output points).

```
IF ( nLev > nThinLev*2 ) THEN
  skip1 = 1 + ( nLev - 1 ) / nThinLev
  skip2 = 1 + nLev
  nSamp = 1 + ( nLev - 1 ) / skip1
END IF
```

If the number of input levels exceeds the maximum requested number of output thinned levels, but is less than twice the maximum, the algorithm rejects 1 in N levels (giving a total of nSamp output points).

```
IF ( nLev > nThinLev .AND. nLev < nThinLev*2 ) THEN
  skip1 = 1
  skip2 = MAX ( nLev / ( nLev - nThinLev ), 2 )
  nSamp = nLev - nLev / skip2
END IF
```

If the number of input levels is less than the maximum requested number of output thinned levels, no thinning occurs (skip1=1, skip2 = nLev+1).

2.4 Thinning by interpolation ropp_io_thin_fixed

Thinning methods requiring interpolation (LIN, LOG, SGLIN, SGLOG, ASGLIN, ASGLOG) are implemented by calling routine ropp_io_thin_fixed.

```
CALL ropp_io_thin_fixed ( nLev,      FullLev, FullVal, &
                        nThinLev, ThinLev, ThinVal, &
                        Method,   DEBUG,   sigma )
```

This routine interpolates input data values to a fixed (thinned) number of levels. For LOG interpolations, if either of the pair of values being interpolated are negative, linear interpolation is performed for that fixed height point. If the full resolution input height range does not fill the range of thinned fixed height, the out-of-range thinned data are set to missing values.

2.4.1 Pre-smoothing

If methods SGLOG, SGLIN, ASGLOG or ASGLIN are selected, the data are first smoothed using Savitzky–Golay (S–G) filtering (3). This provides a local smoothing without introducing significant correlations (1). The standard S–G method applies a kernel filter with linear weighting over a window of ± 1 point centred on each original data point. The kernel filter is data independent.

```
CALL ropp_io_thin_sg ( nLev, Val )
```

For the ASGLOG and ASGLIN cases, an adaptive S–G method is applied. This performs more smoothing, where the smoothing kernel window is

```
nsmooth = MAX [ 2, thinres / fullres ]
```

where `thinres` is the resolution (i.e. vertical level separation distance) of the output thinned data levels and `fullres` is the resolution of the input data levels. The filter is designed such that vertical correlations remain insignificant, as the smoothing window over the original data is always less than the output resolution. If the ratio of the original levels to the thinned level resolutions is smaller than 4:1, the adaptive S–G is equivalent to the standard S–G.

```
CALL ropp_io_thin_sg ( nLev, Val, npoints=INT(nsmooth/2), order=2 )
```

Optional keyword `npoints` specifies the number of points left and right of centre for applying the smoothing kernel and `order` specifies the order of the fitted polynomial.

2.4.2 Interpolation

Following the (optional) pre-smoothing, profile data are interpolated from their original levels to the specified (thinned) output levels.

- Input data are sorted by level before interpolating (i.e. input data are sorted into a monotonic function of height).
- The interpolation uses just the points immediately above and below each output level (source values).
- Interpolation may be done linearly or in log space.
- Output thinned data at a particular target fixed level are set to missing if either
 - the fixed level falls outside the height range covered by the input profile,
 - one of the source values above or below the output level is missing,
 - one of the source level values above or below the output level is missing
 - both source levels are the same (or the gap between levels is less than 1 mm),
 - the gap between source levels is larger than the interval between fixed thinned levels.
- For LOG interpolation, if either source value is less than zero, linear interpolation is performed.

Note that only those parameters which vary logarithmically with height — *viz.* bending angles, refractivity, pressure and humidity — are interpolated thus if one of the ASGLOG, SGLOG or LOG methods are selected; all other parameters are interpolated linearly as if ASGLIN, SGLIN or LIN methods respectively, had been requested. Otherwise the same method is applied to all parameters in all profiles.

2.5 Limitations

- If there is no Level 2a profile present, the fixed Impact Height levels from the control file cannot be converted to geometric and hence geopotential heights, so Level 2b data cannot be thinned.
- The same thinning method is applied to all profiles (except for ASGLOG/SGLOG/LOG methods for which logarithmic interpolation is only applied to bending angle, refractivity, pressure and humidity profiles, and ASGLIN/SGLIN/LIN for all other parameters)
- The fixed-level thinned heights may be output in terms of Impact Parameter (Level 1b), geometric (Level 2a) or geopotential (Level 2b) heights but all are on the same physical levels as defined in the thinning control file. The geometric and geopotential heights will therefore not have fixed *values* from profile to profile

-
- For the SAMPLE method, the number of output levels will vary according to the number of original levels; in general, there will be fewer than the specified maximum.
 - As it is currently implemented, transforming latitude to 0-360 degree range before thinning will lead to errors in the interpolation of points which span 0 degrees longitude (Greenwich meridian) or 170 degrees, or directly go over the North or South pole. A 'spike' may appear in the longitude value at the height at which the profile crosses the critical geographic point.
 - GPS-LEO ray azimuth angle values spanning due North could be similarly affected by the interpolation at the height of the critical crossing point 0 degrees with respect to North, since currently no account is taken for interpolating points spanning 0 and 360 degrees. A revised approach for thinning latitude, longitude and azimuth simultaneously may be implemented for future ROPP versions.

Bibliography

- [1] GRAS SAF. Mono-dimensional data thinning for GPS Radio Occultations. SAF/GRAS/METO/REP/GSR/001, 2007.
- [2] GRAS SAF. The Radio Occultation Processing Package (ROPP) User Guide. Part I: Input/Output module. SAF/GRAS/METO/UG/ROPP/003, Version 3.0, 1 March 2009.
- [3] W.H Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in Fortran – The art of scientific computing*. Cambridge University Press, Cambridge, New York, 2nd edition, 1992.

A Thinning Levels

The list of impact heights used for thinning the data with ROPP SG is shown below. This list has been defined by C. Marquardt, using the criteria of having a maximum of four points per Fresnel diameter. This set is the set of thinned bending angle fixed levels used for generating GRAS SAF BUFR products.

2094.240	2211.910	2330.080	2449.710	2569.340	2690.920
2812.500	2935.550	3058.110	3183.110	3308.110	3433.590
3560.550	3687.500	3815.920	3944.340	4073.240	4203.120
4333.980	4465.330	4597.170	4729.980	4862.790	4997.070
5131.350	5266.110	5401.860	5538.090	5675.290	5812.990
5951.170	6090.330	6229.980	6370.120	6510.740	6652.830
6794.920	6937.990	7082.520	7226.070	7371.580	7517.090
7664.060	7811.040	7958.980	8107.910	8257.320	8407.230
8559.080	8710.940	8863.280	9017.090	9171.390	9326.660
9483.400	9639.650	9797.360	9956.050	10116.70	10276.90
10438.50	10600.60	10764.20	10927.70	11093.30	11260.30
11426.80	11595.20	11765.10	11935.50	12106.90	12279.80
12453.60	12627.90	12804.20	12981.00	13159.20	13338.90
13519.50	13701.70	13884.30	14068.80	14254.90	14441.90
14629.40	14818.40	15010.30	15201.20	15395.50	15590.80
15786.60	15984.40	16182.60	16383.80	16585.90	16789.10
16994.60	17200.70	17408.20	17617.70	17828.60	18041.00
18255.40	18470.70	18689.00	18907.70	19128.90	19350.10
19574.20	19800.30	20027.30	20256.80	20486.80	20718.80
20953.10	21189.50	21426.30	21665.00	21905.80	22149.40
22393.60	22639.20	22887.70	23136.70	23388.70	23642.10
23897.00	24153.30	24412.10	24671.90	24933.10	25196.80
25462.40	25728.50	25997.60	26267.10	26539.10	26813.00
27088.40	27364.70	27643.10	27922.40	28203.60	28486.30
28770.00	29056.60	29342.80	29631.80	29921.40	30212.40
30503.90	30797.40	31092.80	31387.70	31684.60	31982.40
32281.20	32581.10	32881.30	33182.60	33484.90	33787.60
34091.30	34396.00	34700.70	35005.90	35311.00	35617.70
35923.30	36230.00	36537.60	36843.80	37150.90	37458.50
37765.10	38072.30	38379.40	38687.00	38994.60	39301.30
39608.40	39915.50	40223.10	40530.80	40836.90	41144.50
41452.10	41758.80	42065.90	42373.50	42681.20	42988.30
43295.40	43603.00	43910.60	44218.30	44524.40	44832.00
45139.60	45447.30	45754.40	46061.50	46369.10	46676.80
46983.90	47291.50	47599.10	47906.70	48214.40	48520.50
48828.10	49135.70	49443.40	49750.50	50058.10	50365.70
50673.30	50980.50	51288.10	51595.70	51903.30	52210.00
52517.60	52825.20	53132.80	53439.90	53747.60	54055.20
54362.80	54669.90	54977.50	55285.20	55592.80	55899.90
56207.50	56515.10	56822.80	57129.90	57437.50	57745.10
58052.70	58360.40	58667.50	58975.10	59282.70	59590.30
59896.50	/	/	/	/	/

GRAS SAF Reports

SAF/GRAS/METO/REP/GSR/001	Mono-dimensional thinning for GPS Radio Occultation
SAF/GRAS/METO/REP/GSR/002	Geodesy calculations in ROPP
SAF/GRAS/METO/REP/GSR/003	ROPP minimiser - minROPP
SAF/GRAS/METO/REP/GSR/004	Error function calculation in ROPP
SAF/GRAS/METO/REP/GSR/005	Refractivity calculations in ROPP
SAF/GRAS/METO/REP/GSR/006	Levenberg-Marquardt minimisation in ROPP
SAF/GRAS/METO/REP/GSR/007	Abel integral calculations in ROPP
SAF/GRAS/METO/REP/GSR/008	ROPP thinner algorithm
SAF/GRAS/METO/REP/GSR/009	Refractivity coefficients used in the assimilation of GPS radio occultation measurements

GRAS SAF Reports are accessible via the GRAS SAF website <http://www.grassaf.org>.