TN15.2

Aeolus observational requirements and the measurement grouping algorithm

Authors

MICHAEL RENNIE
ECMWF

JOS DE KLOE
KNMI

CHANGE LOG

Version	Date	Comment
1.0	5 Feb 2013	MR: Initial attempt
1.1	15 Feb 2013	Comments from LI
1.2	19 Feb 2013	Comments from JdK
2.0	21 Nov 2013	Comments from AGS

Contents

1 Introduction

This technical note starts by recalling the user requirements for the horizontal resolution and sampling of wind observations as was suggested in TN15.1 (Inventory of Aeolus Target Assimilation Systems, [RD1]) and by the Aeolus Impact Studies (see e.g. [RD5]).

Given the move of Aeolus to continuous mode operation and the consequent possibilities for measurement accumulation strategies, the L2B team have developed an algorithm to group together L1B measurements in such a way that the user requirements can be met. The "Grouping Algorithm", as it is referred to, has now been implemented in the first CM L2B processor (v2.00), which was delivered in December 2012.

The Grouping Algorithm is described and testing with a variety of scenarios is reported to demonstrate the expected behaviour. A brief section about the new L2B/C file format is then provided. The final section provides suggestions of improvements that could be made to the Grouping Algorithm in future.

This technical note was requested by ESA for Sub-task 1.2 of [AD1] and further documented in [AD2], work package 2720.

1.1 Documents

1.1.1 Applicable documents

	Title	Ref	Ver.	Date
[AD1]	Statement of Work for Change Request #4 "Aeolus Level 2B/C Processor - Implementation of Continuous Mode Operations & Extended Pre-Launch Support". Contract 18555/04/NL/MM	AE-SW-ESA-GS-038	1.0	Dec 2010
[AD2]	ESTEC Contract No. 18555/04/NL/MM "Change request for CCN No.4"	N/A	1.0	Feb 2011

1.1.2 Reference documents

	Title	Ref	Ver.	Date
[RD1]	TN15.1 Inventory of Aeolus Target Assimilation Systems	AE-TN-ECMWF-GS-151	1.3	Oct 2012
[RD2]	ADM-Aeolus level-2B algorithm theoretical baseline document	AE-TN-ECMWF-L2BP-0024	2.4	Dec 2012
[RD3]	The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study	By Z. Liu and F. Rabier, Q. J. R. Meteorol. Soc. (2002), 128, pp. 1367-1386		2002
[RD4]	TN3.1a Test cases for the L2B processor	AE-TN-KNMI-GS-0031a	1.0	Feb 2011
[RD5]	TN3 of ESA study contract 4000104080: Synthesizing of draft Aeolus observation requirements, collection of simulated observations and support to VAMP CCN2 contract studies	AE-TN-ECMWF-impact-study-003	4.1	Jan 2013
[RD6]	ADM-Aeolus Level-2B/2C Processor Input/Output Data Denitions Interface Control Document	AE-IF-ECMWF-L2BP-001_20121211_IODD_Iss 2.00	2.00	Dec 2012

1.2 Acronyms

ATBD	Algorithm Theoretical Baseline document
BM	Burst mode
BRC	Basic Repeat Cycle
CM	Continuous mode
DA	Data assimilation
DEM	Digital Elevation Model
DWL	Doppler Wind Lidar
ECMWF	European Centre for Medium-Range Weather Forecasts
EGM	Earth Gravitational Model
HLOS	Horizontal Line Of Sight
IODD	Processor Input/Output Data Denitions Interface Control Document
KNMI	Royal Netherlands Meteorological Institute
L1B	Level-1B
L2B	Level-2B
L2Bp	L2B processor
N/A	Not applicable
NWP	Numerical weather prediction
QC	Quality control
RMA	Reference model atmosphere
SNR	Signal to noise ratio
SRD	System requirements document
TBD	To be determined
TN	Technical note
VHAMP	Vertical and Horizontal Aeolus Measurement Positioning
WGS	World Geodetic System
XML	Extensible Markup Language
ZWC	Zero wind correction

2 User requirements

The technical note 15.1 (Inventory of Aeolus Target Assimilation Systems, [RD1]) provided some insight into what the potential users of Aeolus L2B wind data require in terms of observation horizontal resolution. This information was useful to the L2B/C team when defining new requirements for the L2B processing which are suitable for the continuous mode (CM) operation. The Aeolus scientific impact studies also provided suggestions of a sensible averaging length-scale for Aeolus CM L2B winds. Other limitations/requirements on the L2B processing are imposed by the L1B data itself, therefore a compromise between what users would like and what Aeolus data will allow has been sought.

The main message from TN15.1 was that users have a wide range of preferences for horizontal resolution of Aeolus wind observations. Those with high resolution (mesoscale) models felt that sub-BRC scale (less than 86 km) wind observations would be most useful (despite being informed about the consequent increase in noise); whilst others with relatively coarse resolution global models thought that super-BRC scale (100-150 km) resolution would be best. Within the global model community (with models of similar grid sizes), estimates of the model effective resolution (and hence a sensible representative scale for the wind) varied widely e.g. ECMWF thought that 60-80 km was an appropriate estimate for a 2015 model resolution, whereas Environment Canada thought 150 km.

The Aeolus scientific impact study by ECMWF, [RD5], recommended that horizontal averaging of less than one BRC (<86 km) would most likely be appropriate for ECMWF given the model resolution planned for 2015. Prior to the real data becoming available, it is very difficult to determine what the optimal horizontal resolution is for Aeolus in a particular NWP system, since it depends on the interaction between model resolution, observation resolution, accuracy and density in a rather complex way (e.g. see [RD3]).

Given the above suggestions, the recommendation regarding the L2B CM processing is:

To implement a flexible measurement grouping algorithm, allowing wind observations to be constructed from the measurement horizontal-scale up to any number of measurements which can be greater than the BRC horizontal-scale. It should also be able to run with a single-BRC grouping to replicate the results of the BM L2B processor.

The latter suggestion, regarding being able to run with single BRCs, like for BM, is so that scientific comparisons of CM results with BM results can be made, to ensure the algorithm changes have not introduced any bugs to the code. With ESA's announcement that the laser output energy will likely be reduced to 80 mJ (at least for the start of the mission), the measurements will require averaging to greater than the BRC to retrieve winds of 2 m/s random error for the Rayleigh channel.

3 The Grouping Algorithm

After excluding some initial suggestions, e.g. concatenating L1B BRCs into larger so called super-BRCs via an intermediate file (in L1B format), a plan was produced by KNMI to implement a "Grouping Algorithm" as part of the L2B processing. This algorithm has now been implemented by KNMI, and was part of the L2BP v2.00 (released in December 2012).

The main aim of the algorithm is to allow measurements to be grouped along the orbit independently of the BRC definition, therefore providing flexible horizontal wind resolution as required by users.

The next section which describes the Grouping Algorithm has been based on the description in the ATBD [RD2].

3.1 Overview of the measurement Grouping Algorithm

Firstly we provide some definitions of Aeolus data to help describe the Grouping Algorithm, which is illustrated in Figure 1 (this shows a vertical plane of Aeolus measurement-bins). One can see how the smallest constituent, the measurement-bin, is part of a measurement, and applies to a range-bin, and can be part of what we refer to as a group of measurements.

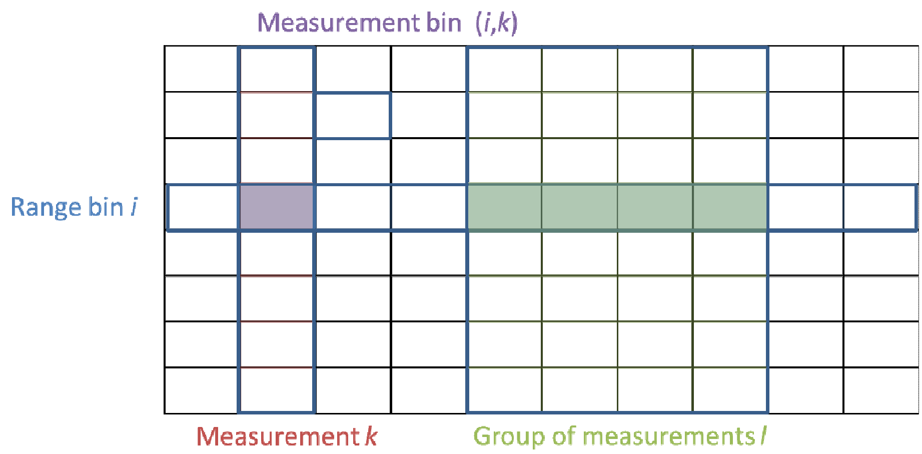


Figure 1. How Aeolus data is defined in terms of measurements, range-bins, measurement-bins and groups of measurements.

The L2B processor produces observations of HLOS wind component, which we call wind results, to reflect the fact that they are not profiles of winds, but single wind observations. The observations are the output of wind retrievals performed on averages of Aeolus measurement-bin scale (2.85 km) data. Averaging of measurement-bin data produces observations more suitable for NWP in terms of noise and representative scale.

This section explains how the L2B processor groups together measurements. After grouping, the processing proceeds to classify the measurement-bins by type, followed by a wind retrieval on those classified measurement-bins to produce a wind observation.

As explained earlier, the aim of the Grouping Algorithm is to improve the use of continuous mode (CM) Aeolus data, in line with the requirements of potential Aeolus NWP users. In CM there is no reason for the wind retrieval to be restricted to using the measurement-bins from only one BRC. With BM, averaging within only one BRC made sense, since the 150 km horizontal gaps between BRCs were too large to consider averaging across; this is illustrated in Figure 2. Note that for CM the BRCs are 86 km, whereas they were 50 km for the BM configuration.

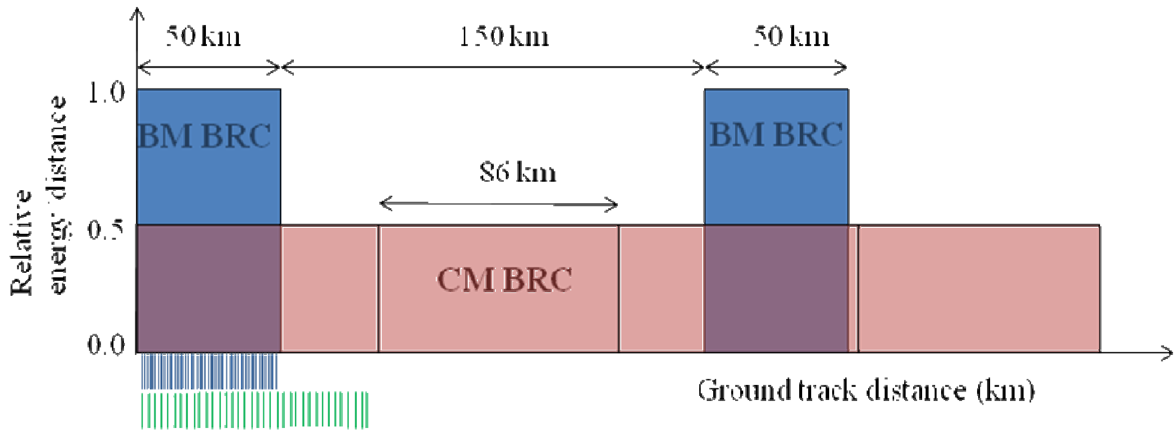


Figure 2. Comparison of continuous mode (CM) operation to burst mode (BM) operation in terms of energy delivered to the atmosphere as a function of along-track distance.

The Grouping Algorithm forms groups of adjacent measurements independently of the defined BRCs. The groups are chosen separately for each channel (Rayleigh and Mie); since each channel has its own properties in terms of signal quality and range-bin definitions.

The grouping can be thought of as creating variable sized BRCs i.e. 2D arrays of measurement-bins, with horizontal dimension equal to the new number of measurements in a group, and vertical dimension the 24 vertical range-bins — it does no vertical regridding. **A group can have a horizontal extent from one measurement up to all the measurements in the L1B file if required**. The grouping starts from the first (earliest) measurements in the L1B file — constructing groups, one after another, along the orbit. A minor consequence is that rather small groups could be formed where grouping is made to stop (i.e. a threshold is exceeded (see below) or at the end of a sequence of L1B data), hence the groups can vary in size along the orbit; there is no minimum size restriction at present, only the maximum size restriction specified by the user.

Another feature is that the chosen groups are dependent on where within a sequence of Aeolus data the grouping algorithm is started, e.g. if a L1B file was split into two smaller files, the grouping in the second smaller L1B file could be different to that chosen with the original larger file. This (probably) prohibits parallel processing of L1B data e.g. if the L2B processor was done via a subroutine version of an NWP system (unless the one-BRC option is chosen, see “Classic” below). This is not considered a problem, because the L2B processor is fast.

The horizontal extent (and therefore the number of measurements) in a given group is determined by the parameters specified in the AUX_PAR_2B file and how these interact with a given sequence of L1B data. Note that it is possible to select the group size to be equal to one BRC, in which case the processing is similar to the old burst-mode processing (see “Classic” below).

Inputs to the Grouping Algorithm:

- L1B geolocation information
- Grouping parameters from the AUX_PAR_2B_file

Outputs from grouping algorithm:

- The resulting indices defining which measurements belong to which group

The algorithm:

The grouping algorithm method is chosen by settings in the AUX_PAR_2B file. The chosen method determines the group size and hence an upper limit on horizontal observation size after measurement selection. There are two methods available:

- **Classic:** this mimics the old Burst Mode method, in that it creates groups which always exactly match the BRCs defined in the L1B product.
- **Advanced:** this takes thresholds (defined below) specified in the AUX_PAR_2B file, and tries to construct groups as large as possible, given the thresholds, within the available set of measurements in a single L1B product file. The chosen threshold values allow control of which measurements will be averaged when creating observations.

  The Advanced method is controlled by the following parameters (independent parameters for Mie and Rayleigh channels):

  - **Max_Vertical_Rangebin_Misalignment_{Mie,Rayleigh}**: This parameter stops inappropriate averaging if measurement range-bins are vertically shifted relative to one another by more than a tolerated value (Figure 5). The check is hence done separately on the Rayleigh range bins for the Rayleigh observation averaging and on the Mie range bins for the Mie observation averaging. The tolerated misalignment value is the maximum allowed vertical altitude difference between {Mie, Rayleigh} measurement range-bins with the same range-bin index. If a measurement range-bin is found to have a larger altitude difference to the rest of the group, a new group will be started. Range-bin altitudes (referenced to the geoid) can change due to steps in the on-board Digital Elevation Model and the geoid undulation (relative to the ellipsoid reference frame that is used by the L1BP) which are used to command the altitude (range) of the lowest range bin and due to changes in the commanding of the vertical range-bin distribution along the orbit.
  - **Max_Horizontal_Accumulation_Length_{Mie, Rayleigh}**: This parameter allows control of the maximum horizontal-scale of an observation (Figure 3 and 4). It is the maximum horizontal distance between the first and last {Mie, Rayleigh} measurements in a group. If a measurement is found at a larger distance, then a new group will be started. Increasing the accumulation length includes more measurements in the group and hence also reduces the noise in the resultant winds, but at the cost of horizontal resolution, which suppresses the detection of atmospheric variability. Note that the final horizontal scale of a wind result also depends on the cloud/aerosol dependent classification algorithm (see the ATBD [RD2] for further details). Mie winds will most likely require shorter averaging lengths than Rayleigh to achieve a certain SNR due to the stronger backscatter from particles than molecules.
  - **Max_Allowed_Gap_Between_{Mie, Rayleigh}_Measurements**: This parameter controls the maximum distance of missing {Mie, Rayleigh} measurements before a group definition is closed and a new group is started. This ensures that large gaps between the data in a group can be eliminated i.e. we do not want data only at either end of a large group, with nothing in between to produce a wind observation. With CM data, there should normally not be many gaps in the data along the orbit, but perhaps this could happen due to e.g. calibration modes, or transmission or storage malfunctions.

In Figure 3 a "measurement map" illustrates how the Grouping Algorithm works showing the location of measurement-bins in the vertical and horizontal (along-track) dimensions. Each small rectangle represents a measurement-bin (similar to Figure 1). Note that the vertical range-bins can vary in size, typically from 250 m to 2 km. The resulting groups of the algorithm are indicated by green arrows (at the bottom), showing the horizontal extent of the groups. In this example, the grouping algorithm has been set such that groups are formed from five adjacent measurements (horizontally), i.e. Max_Horizontal_Accumulation_Length = 5*measurement_width (km)

Measurements within a group are later assigned to specific observations (see ATBD [RD2]) — in Figure 3 the observation index to which a measurement range bin belongs is also shown (the number in each measurement-bin). This scenario would be typical of Rayleigh-clear observations, in a clear atmosphere, since all measurement-bins in a group on a given range-bin could be used. The light blue shading highlights the measurements that go into observations numbers 1 and 54.

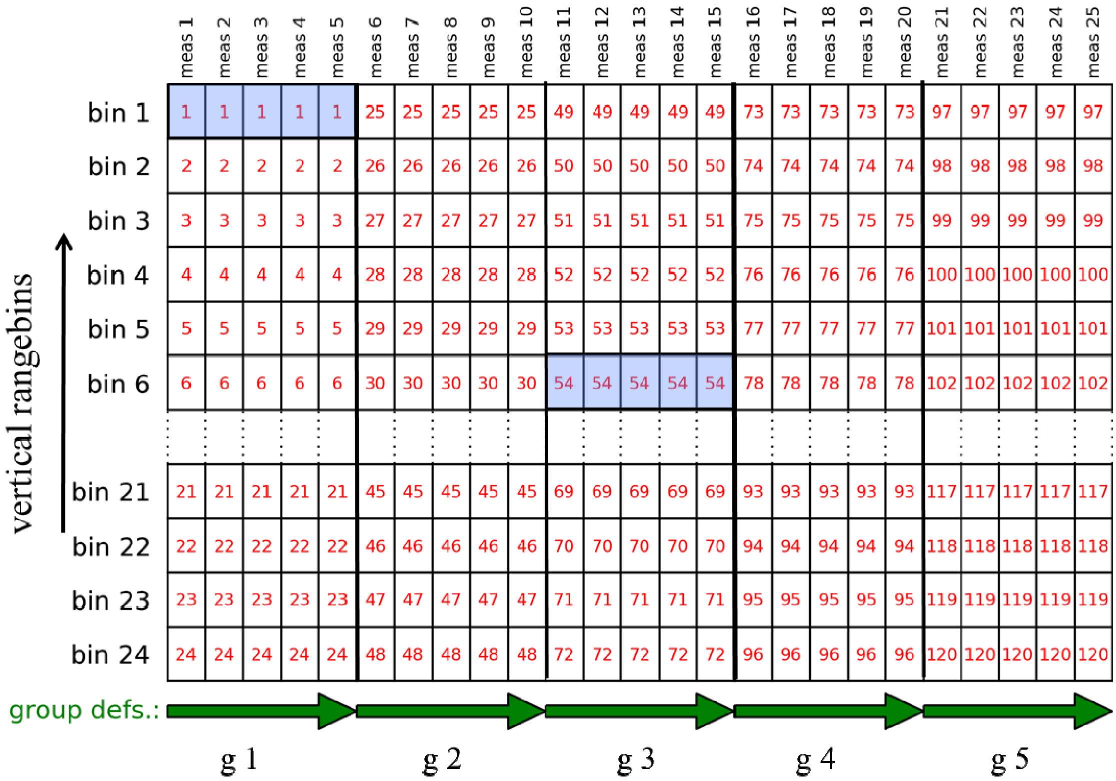


Figure 3. A "measurement map" diagram to illustrate how the Grouping Algorithm works. The vertical axis represents the 24 vertical range-bins of Aeolus and the horizontal axis represents either time or distance along the orbit track (or measurement number). The green arrows at the bottom indicate the number of measurements (and hence horizontal distance) over which groups extend. The number within each measurement-bin shows to which wind observation it has been assigned.

Figure 4 (similar to Figure 3) illustrates how the grouping is not required to stop if the threshold in the "Advanced" settings are not reached. Only one group is produced in this example, with all the measurement-bins on a given range-bin being assigned to one observation.

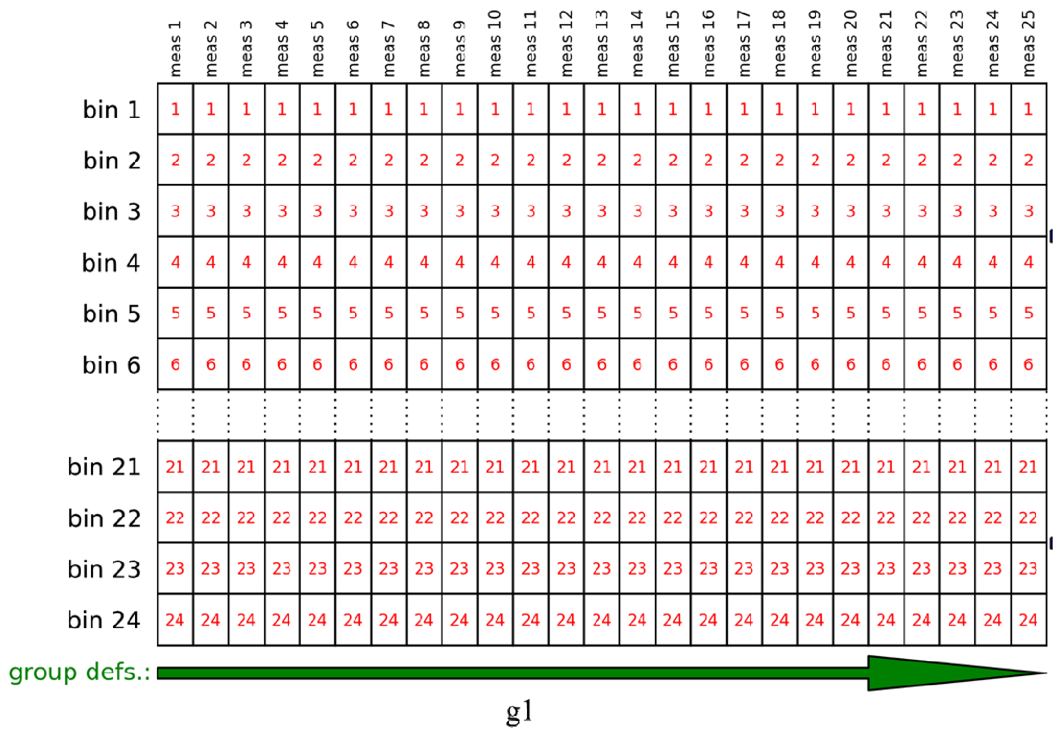


Figure 4. A "measurement map" illustrating a case when the Grouping Algorithm thresholds have not been reached, such that one group has been formed from all the available measurements.

Figure 5 illustrates how groups can start/end if the vertical displacement of range-bin heights has exceeded the defined threshold (i.e. Max_Vertical_Rangebin_Misalignment). Here the observations could be for Mie-cloudy observations, since they are not on every range-bin, or the QC has not allowed the other measurement-bins to be used in the wind retrieval. To be clear here we show the measurement-bins used to produce observations following the grouping algorithm and then followed by scene classification.

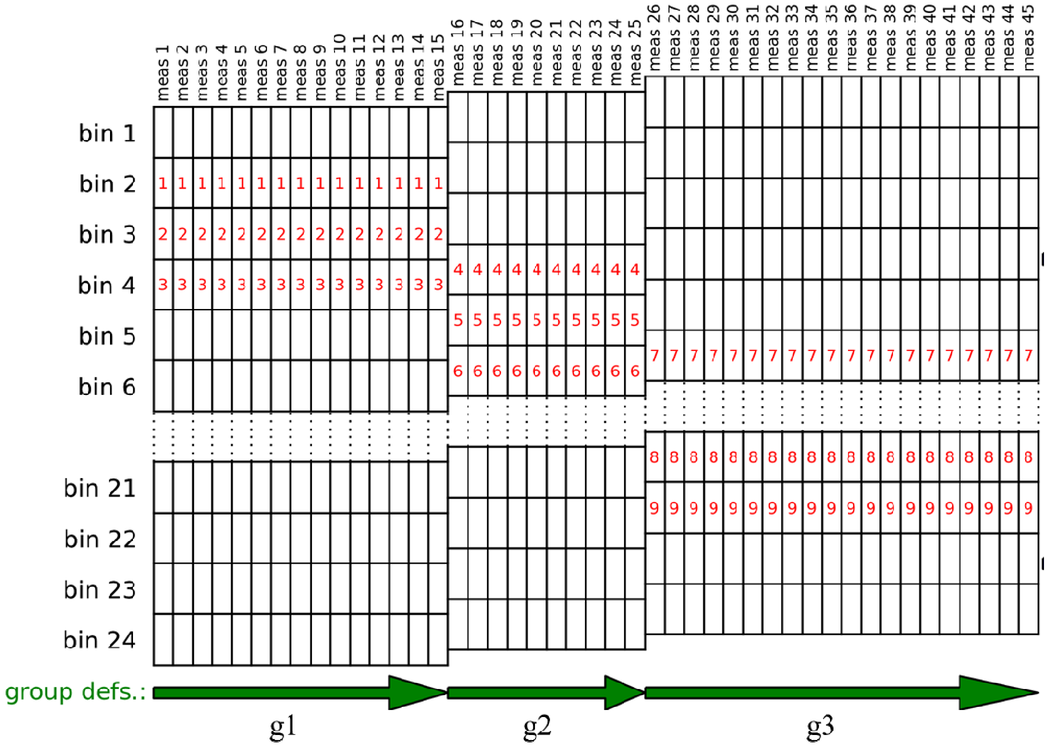


Figure 5. A "measurement map" illustrating the effect of changes of range-bin height definition upon the grouping. Here only some measurement-bins are assigned to observations, which is possible due to scene classification or some form of QC based on signal strength (see ATBD [RD2]). This figure hence shows the groups after the successive scene classification and QC steps.

Further detailed information on the L2B processing steps taking place after the groups are formed, e.g. classification, averaging and the wind retrieval, can be found in the ATBD [RD2]. Note the processing after grouping has not changed in the transition from BM to the CM L2B processor.

3.2 Testing of the Grouping Algorithm

This section documents some simple testing of the functionality of the Grouping Algorithm. The test cases are from the CM Chain-of-Processors end-to-end testing (Atm. DB→E2S (v3.02, default settings)→L1B (v6.01, default settings)→L2B (v2.00)). The tests here involve only changing the L2B processing AUX_PAR_2B grouping parameters i.e. the E2S and L1B data remain identical in the tests reported in this subsection. It should be noted that the L1B refined scattering ratio was used by the L2Bp classification. The plotting of results was done via the reading of ASCII dumps of of the L2B products and E2S inputs into IDL (at ECMWF).

3.2.1 Testing the accumulation length criterion

The test case chosen is an academic one, #0027 from the standard TN3.1b testing: 3 cloud layers, 4 BRCs. See [RD4] for further details about the test case. It is a horizontally homogeneous case, with temperature, pressure and aerosol properties from the atmospheric scene single RMA profile mid-latitude winter. A constant input HLOS wind of 50 m/s was taken. In addition successive layers of clouds at 4 km, 9 km, and 16 km altitudes have been artificially added (i.e. not using realistic LITE/CALIPSO attenuated backscatter information). There is no advantage in using realistic cases for functional testing of the Grouping Algorithm.

3.2.1.1 Classic method

This test sets the L2B grouping method to "Classic", ensuring that groups will be formed from one BRC of measurements. That is, in the AUX_PAR_2B file (XML):

```
<BRC_Grouping_Params>
  <Grouping_Method>classic</Grouping_Method>
…
…
</BRC_Grouping_Params>
```

Figure 6 shows the Rayleigh-clear wind results within a vertical plane along the ground track. The horizontal extent of the wind result (as shown in the figure) is calculated from the geolocation information of the first and last measurements selected following grouping and classification. Similarly the vertical extent of the range-bins is shown.

Notice that the maximum horizontal-scale of the wind observations is, as expected, limited to one BRC (the group size chosen) i.e. ~86 km. Due to classification (into clear and cloudy) according to L1B scattering ratio, wind observations can be constructed from measurements which extend over less than the group size (the same behaviour as in the BM L2B processor). This is demonstrated, for example, by the shortened wind observation at ~15 km height in BRC (or group) number 4 (~300 km along-track distance). Note that white colour signifies an absence of wind results. It is shortened due to the presence of cloud (see the Mie results in Figure 7 for an indication of where the cloud is within the plane).

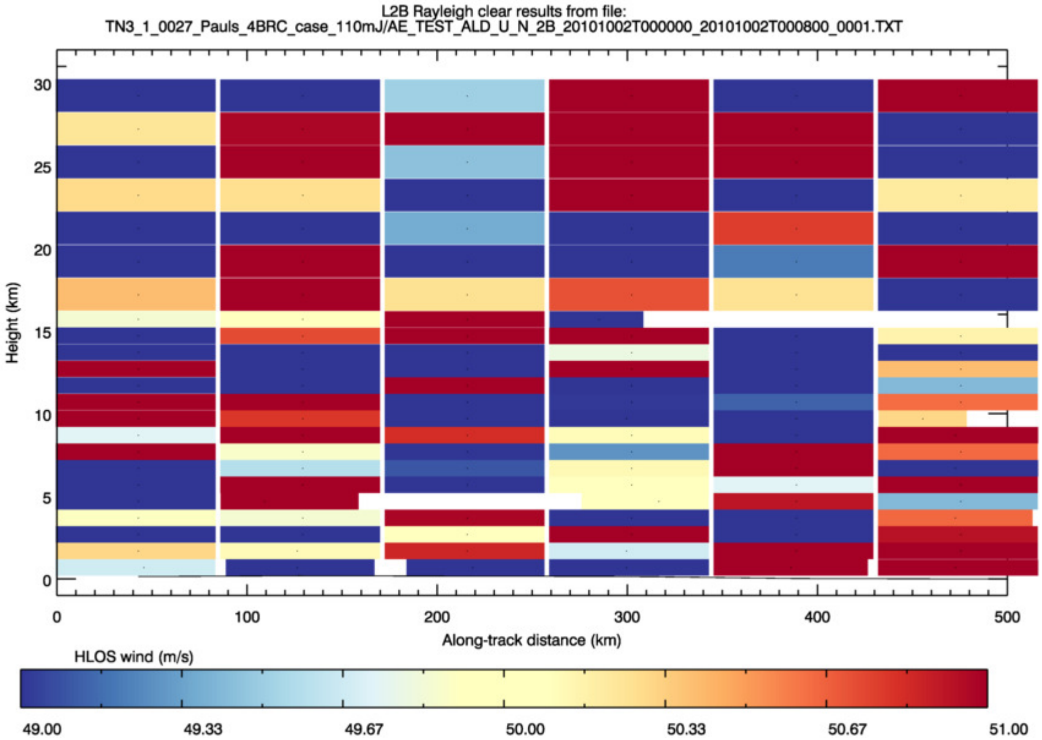


Figure 6. Grouping Algorithm set to "Classic" for Rayleigh-clear results. The plot shows the spatial extent of the wind observations as positioned in the plane along the orbit track. The colours indicate the retrieved HLOS wind (which ideally should be 50 m/s). White signifies an absence of wind results. It can be seen that the maximum horizontal extent of the wind observations is one CM BRC (~86 km).

In Figure 7 it can be seen in that the classification for Mie-cloudy results has produced wind results in areas with particles (in this case the artificial clouds). Given that particle loading can be variable (and the classification is based on noisy backscatter ratio estimates) then it is rarer than in the Rayleigh-clear results to see observations extending over the whole group (1 BRC); but it does occur in this example.

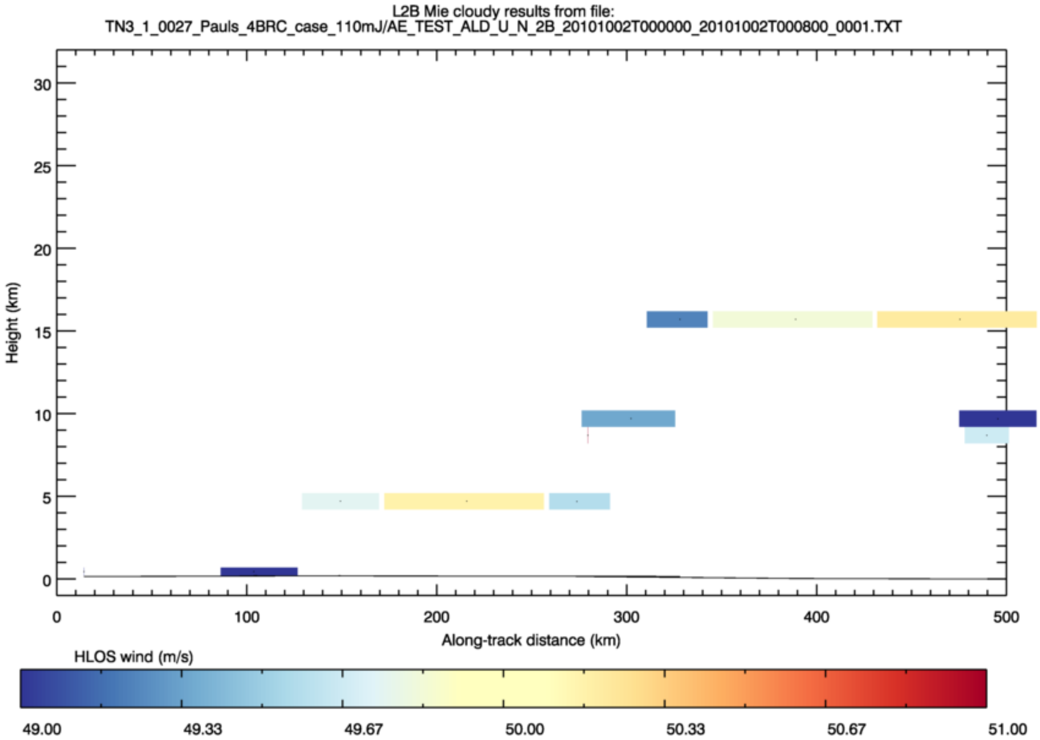


Figure 7. Mie-cloudy results, plotted in the same style as Figure 6. Grouping algorithm set to "Classic".

3.2.1.2 Advanced method set to 200 km

This test sets the grouping method to "advanced", and also sets the maximum horizontal accumulation length to 200 km for both the Mie and the Rayleigh channels. That is, in the AUX_PAR_2B file (XML):

```
<BRC_Grouping_Params>
  <Grouping_Method>advanced</Grouping_Method>
  <Max_Vertical_Rangebin_Misalignment_Mie unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Mie>
  <Max_Vertical_Rangebin_Misalignment_Rayleigh
unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Rayleigh>
  <Max_Horizontal_Accumulation_Length_Mie unit="km">200.0</Max_Horizontal_Accumulation_Length_Mie>
  <Max_Horizontal_Accumulation_Length_Rayleigh
unit="km">200.0</Max_Horizontal_Accumulation_Length_Rayleigh>
  <Max_Allowed_Gap_Between_Mie_Measurements unit="km">160.0</Max_Allowed_Gap_Between_Mie_Measurements>
  <Max_Allowed_Gap_Between_Rayleigh_Measurements
unit="km">160.0</Max_Allowed_Gap_Between_Rayleigh_Measurements>
</BRC_Grouping_Params>
```

In a similar manner to the "Classic" grouping, the wind observations can be constructed from measurements extending up to this maximum horizontal accumulation length, as demonstrated in Figure 8 and Figure 9.

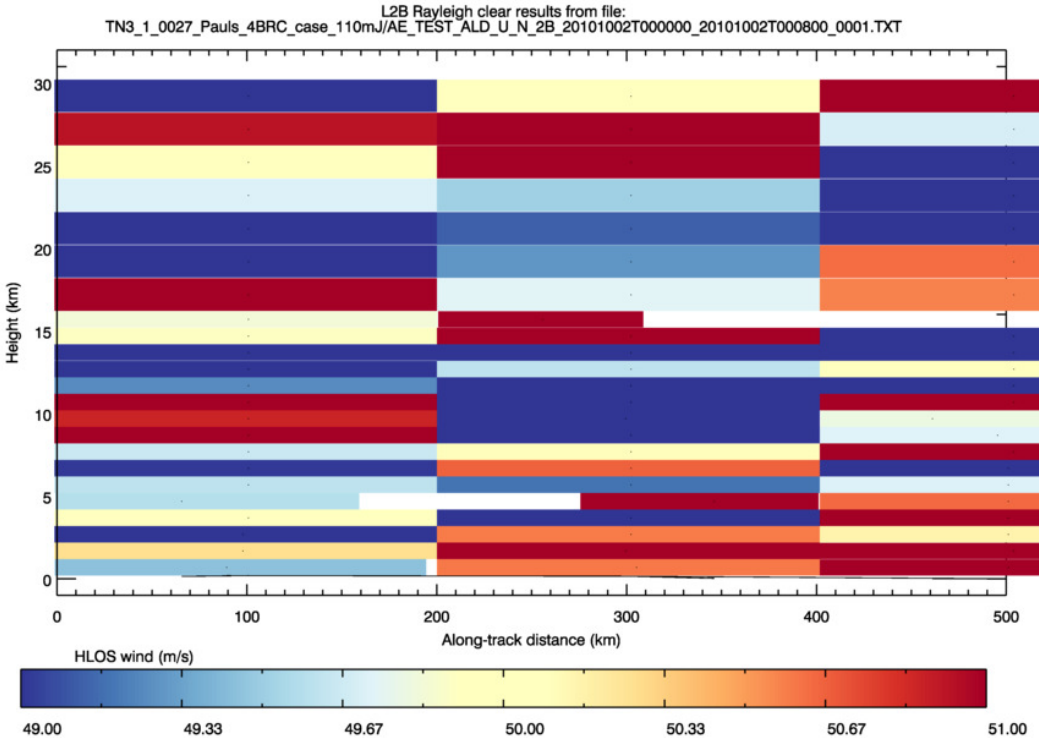


Figure 8. Grouping algorithm set to "Advanced" for Rayleigh-clear wind results. Similar to Figure 6, except the grouping has been set to a maximum accumulation of 200 km. Note the third group continues beyond the plot edge.

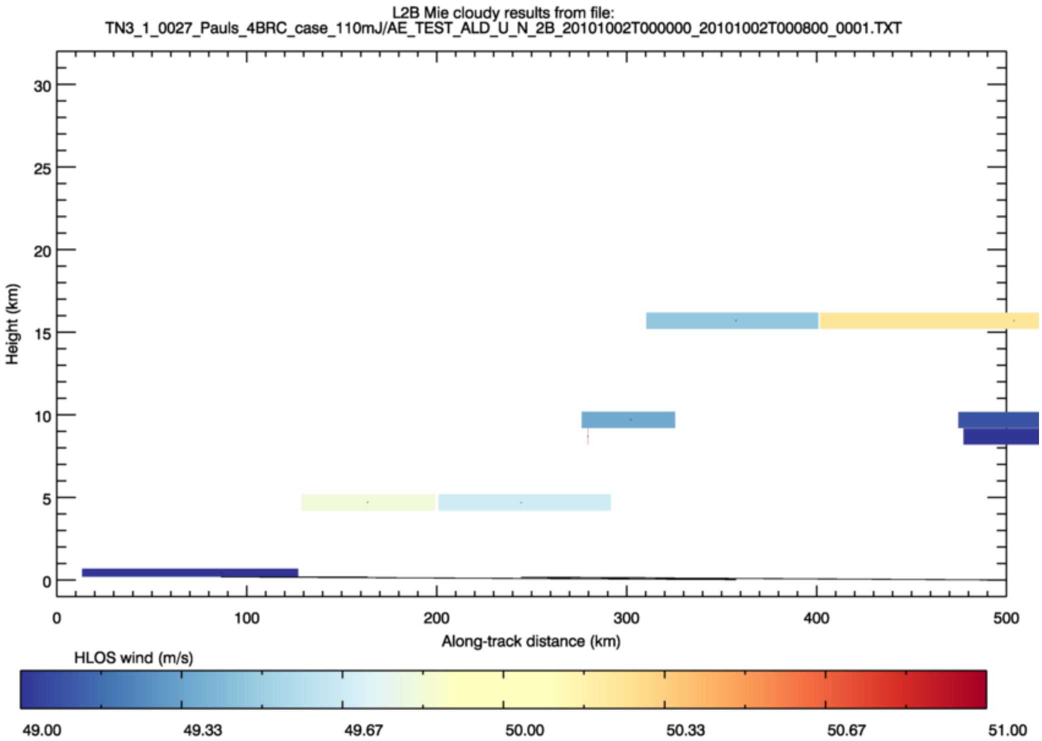


Figure 9. Grouping algorithm set to "Advanced" for Mie-cloudy wind results. Similar to Figure 7, except the grouping has been set to a maximum accumulation of 200 km.

3.2.1.3 Advanced method with different lengths for Mie and Rayleigh

This test is similar to the previous one i.e. "advanced" method, except that the maximum accumulation lengths have been reduced to less than one BRC, and the lengths for Mie and Rayleigh groups have been set to different values: the accumulation length for Rayleigh has been set to 60 km and for Mie to 15 km. That is, in the AUX_PAR_2B file (XML):


```
<BRC_Grouping_Params>
  <Grouping_Method>advanced</Grouping_Method>
  <Max_Vertical_Rangebin_Misalignment_Mie unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Mie>
  <Max_Vertical_Rangebin_Misalignment_Rayleigh
unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Rayleigh>
  <Max_Horizontal_Accumulation_Length_Mie unit="km">15.0</Max_Horizontal_Accumulation_Length_Mie>
  <Max_Horizontal_Accumulation_Length_Rayleigh
unit="km">60.0</Max_Horizontal_Accumulation_Length_Rayleigh>
  <Max_Allowed_Gap_Between_Mie_Measurements unit="km">160.0</Max_Allowed_Gap_Between_Mie_Measurements>
  <Max_Allowed_Gap_Between_Rayleigh_Measurements
unit="km">160.0</Max_Allowed_Gap_Between_Rayleigh_Measurements>
</BRC_Grouping_Params>
```

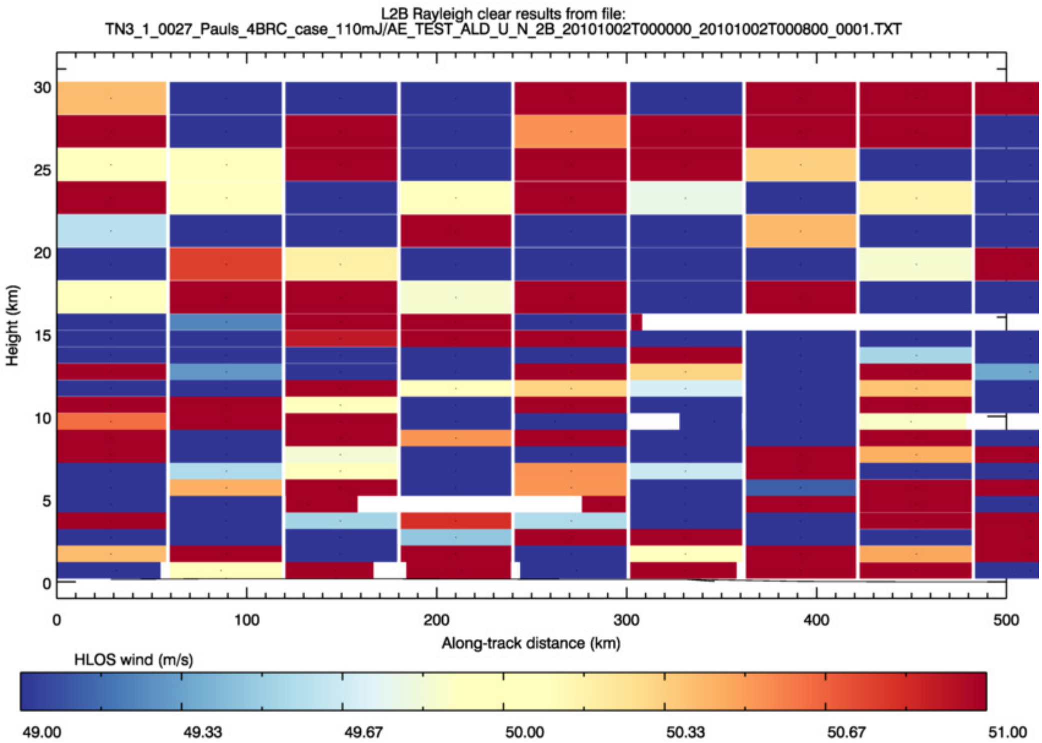



Figure 10. Grouping algorithm set to "Advanced". Rayleigh-clear results with 60 km grouping accumulation length.

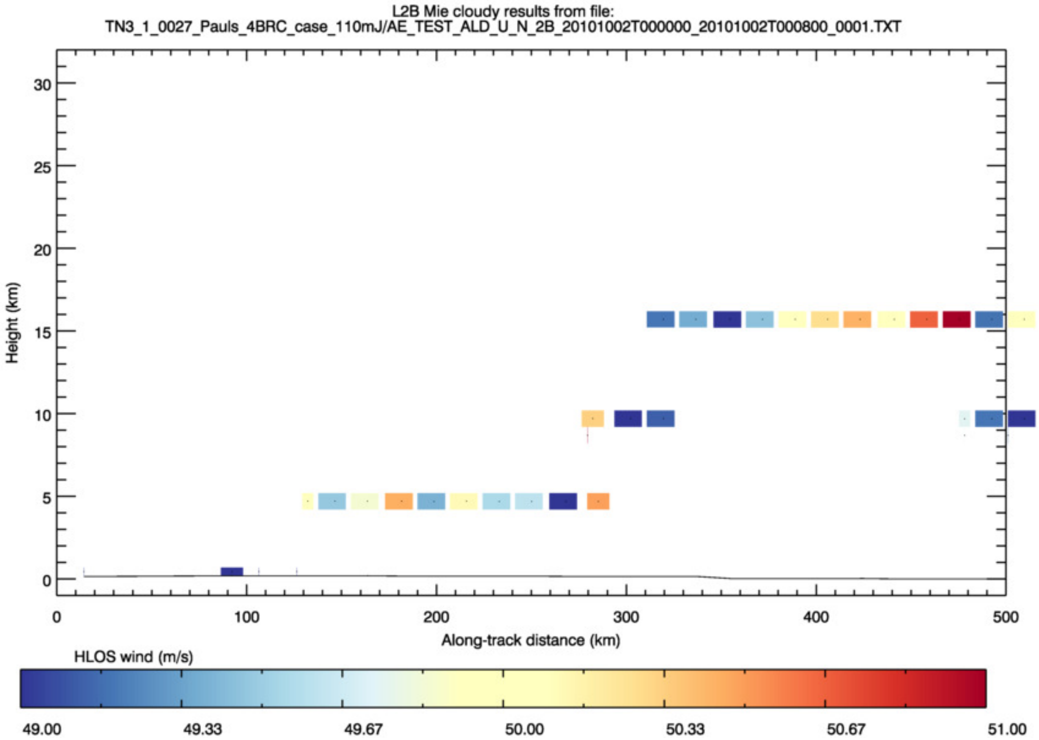


Figure 11. Grouping algorithm set to "Advanced". Mie-cloudy results with 15 km grouping accumulation length.

To demonstrate the nominal quality of the L2B wind results when the Grouping Algorithm is in use, verification of the Rayleigh-clear and Mie-cloudy wind results are shown in Figure 12 and Figure 13 respectively. In particular the plots show statistics of the L2B wind errors (retrieved minus E2S input). This is the same test case as shown in Figure 10 and Figure 11, except the 4 BRCs are repeated 10 times to increase the robustness of the statistics. Note this test case used an E2S laser energy of 110 mJ.

The standard deviation of error for the Rayleigh-clear results is larger than the ESA SRD (which is applicable for 100 km averaging) at about 2.5 m/s in that height range. However, this is expected with the smaller averaging (up to 60 km) which is enforced by the grouping. It can be seen that with only 15 km averaging, the Mie-cloudy results can achieve less than 1 m/s random error, due to the high backscatter from particles. The Mie results apparently show some outliers (just below clouds), so perhaps some tuning of the SNR threshold is required for the Mie-core processing. Note that no QC to remove gross outliers (as part of the verification statistics/plotting) was used in producing the statistics. It is believed that due to the small sample of data, that the outliers are producing a negative mean error (bias).

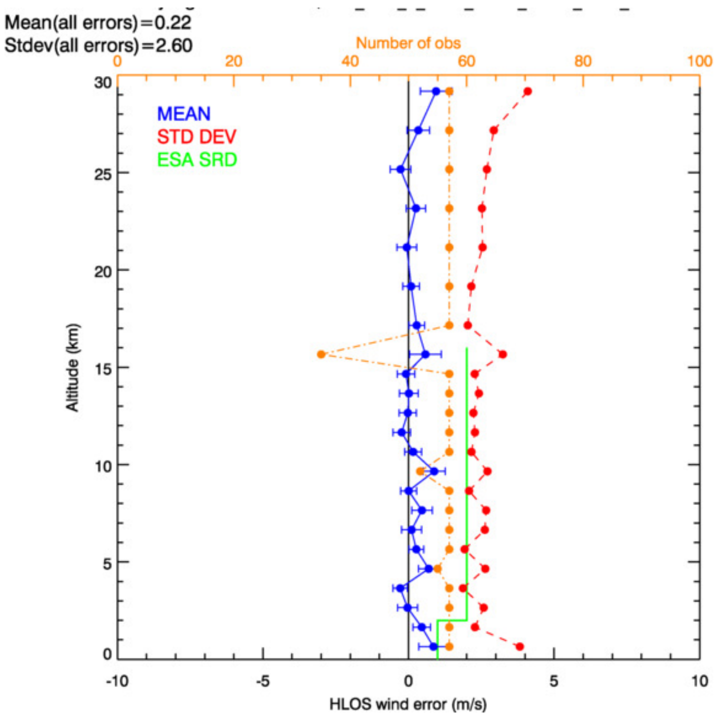


Figure 12. Verification of the L2B wind results against the E2S input winds, Rayleigh-clear, with the "Advanced" grouping set to 60 km. The mean of the errors (retrieved wind minus E2S input) are shown in blue, the standard deviation in red, ESA's System requirements document values for random error are in green and the count of wind results at each range-bin shown in orange.

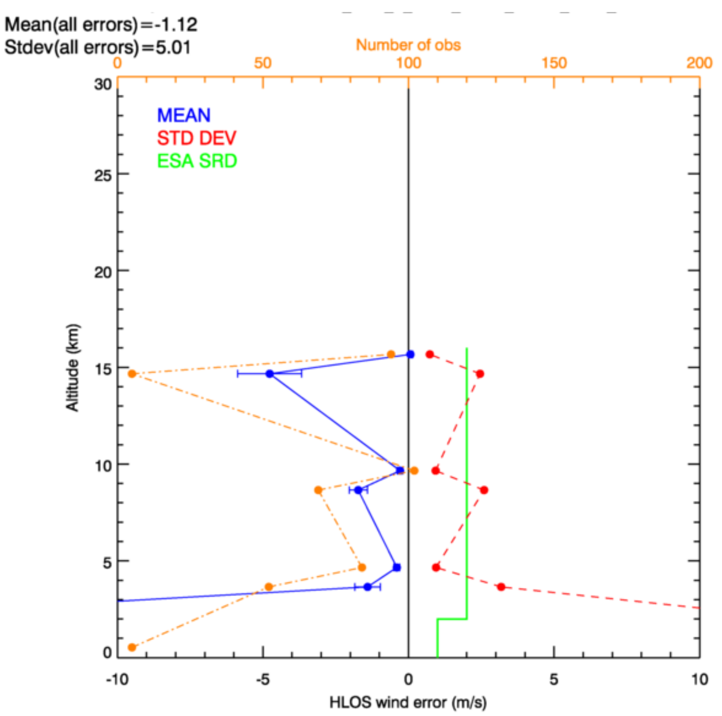


Figure 13. Verification of the L2B wind results against the E2S input winds, Mie-cloudy, with the "Advanced" grouping set to 15 km. The mean of the errors (retrieved wind minus E2S input) are shown in blue, the standard deviation in red, ESA's System requirements document values for random error are in green and the count of wind results at each range-bin shown in orange.

It is clear from the above testing that the Grouping Algorithm is working as expected as regards the maximum accumulation length parameter.

3.2.2 Testing 'edge' effects and the range-bin misalignment criterion

A L1B file will consist of an integral number of BRCs. Only one value of the Digital Elevation Model (DEM) and geoid undulation are reported per BRC (i.e. reported at the BRC level, not at the measurement level). Therefore with non-BRC length groups it is possible to have jumps in range bin altitude definitions within only 1 measurement of the start of the group. That is, if a criterion is triggered part way through the forming of a group, as a BRC boundary has been passed (such as range-bin misalignment), then the group could potentially stop (this of course depends upon the assigned thresholds). The following tests demonstrate this behaviour of forming smaller groups than the chosen maximum accumulation length. Also if the grouping is not set to one BRC, then as the L1B file comes to an end, a shorter end group will be created.

3.2.2.1 Range-bin misalignment

Due to a misunderstanding when doing this testing, it was thought that E2S v3.02 cannot simulate the vertical shifting of range-bins to follow the DEM as will be seen for real Aeolus data (this turns out to be false, it can simulate this). However it was still appropriate to test the effect due to a small variation in the height of the range-bins (which are referenced by the L2BP to height above geoid (EGM96)) as simulated by E2S due to the range-bins being fixed relative to the WGS 84 ellipsoid, and the geoid reference undulating relative to the ellipsoid (the deviations of the EGM96 geoid from the WGS 84 reference ellipsoid range from about −105 m to about +85 m).

Some details:

1. Aeolus L2B winds are provided with heights relative to the geoid (and in particular those heights used by the Grouping Algorithm).

2. The range-bins are fixed relative to the ellipsoid surface (in the current E2S), causing a given range-bin height to vary (as referenced to geoid) due to geoid undulation.

It was noticed that the threshold for maximum vertical range-bin misalignment was triggered with test case #0013, a noise-free case with 8 BRCs, with a cloud around 8-9 km and increasing in altitude along the orbit. See [RD4] for further details about the test case.

The results of test case #0013 can be seen in Figure 14 below. We show the Rayleigh-clear winds for a case with "Advanced" grouping, with the maximum accumulation length set to 360 km and the maximum vertical range-bin misalignment was set to 10 m. That is, in the AUX_PAR_2B file (XML):

```
<BRC_Grouping_Params>
  <Grouping_Method>advanced</Grouping_Method>
  <Max_Vertical_Rangebin_Misalignment_Mie unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Mie>
  <Max_Vertical_Rangebin_Misalignment_Rayleigh
unit="m">10.0</Max_Vertical_Rangebin_Misalignment_Rayleigh>
  <Max_Horizontal_Accumulation_Length_Mie unit="km">360.0</Max_Horizontal_Accumulation_Length_Mie>
  <Max_Horizontal_Accumulation_Length_Rayleigh
unit="km">360.0</Max_Horizontal_Accumulation_Length_Rayleigh>
  <Max_Allowed_Gap_Between_Mie_Measurements unit="km">160.0</Max_Allowed_Gap_Between_Mie_Measurements>
  <Max_Allowed_Gap_Between_Rayleigh_Measurements
```

```
unit="km">160.0</Max_Allowed_Gap_Between_Rayleigh_Measurements>
</BRC_Grouping_Params>
```

One can see that the first group is indeed 360 km long; however the next one is shorter at 240 km, when it potentially could have continued until the end of the data set (the end of the L1B data comes at around 690 km).

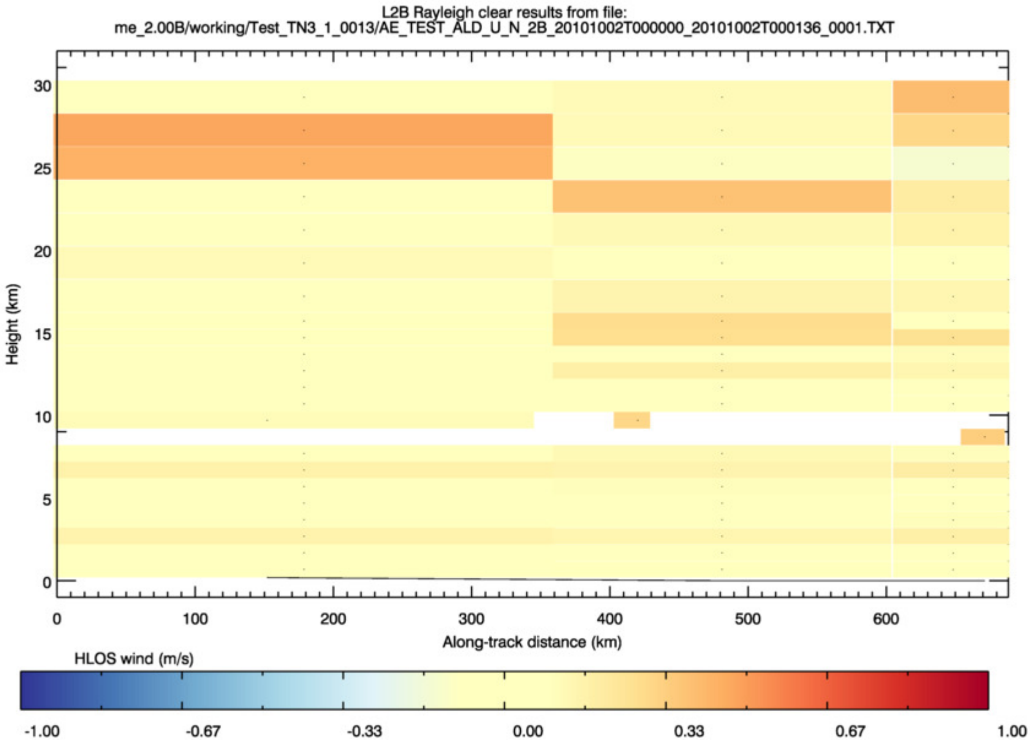


Figure 14. Grouping Algorithm set to "Advanced" for Rayleigh-clear results. The maximum accumulation length was set to 360 km and the maximum vertical misalignment was set to 10 m. It can be seen that the horizontal extent of the wind observations is cut short in the second group.

The standard output of the L2B processor indicates what is happening:

```
RANGEBIN MISALIGNMENT criterion failed for Rayleigh channel: difference
alt_bottom =    10.013218206731835       or

 RANGEBIN MISALIGNMENT criterion failed for Rayleigh channel: difference
alt_top    =    10.041005714243511

merge NOT possible

 meas_groups for the Rayleigh channel:

 Group 1: it is safe to combine overall meas index   1   (BRC/Meas:     1
1 ) upto  125  (BRC/Meas: 5      5 )

 Group 2: it is safe to combine overall meas index   126   (BRC/Meas:   5
6 ) upto   210  (BRC/Meas: 7    30 )

 Group 3: it is safe to combine overall meas index    211  (BRC/Meas:    8
1 ) upto    240  (BRC/Meas: 8   30 )
```

The output indicates the second group was stopped due to the range-bin misalignment criterion being exceeded, since the difference in range-bin altitude between the first measurement of the second group and the 84th measurement exceeded 10 m. This very strict vertical mis-alignment threshold was selected simply to test that processor works as expected. Similar behaviour would be expected if range-bin misalignment was caused by DEM variations or range-bin definition changes.

The example of Figure 14 also demonstrates the "edge" effect in that the final group is cut-off prematurely due to the end of the L1B data; this was only an 8 BRC test case. Obviously these smaller groups will result in noisier wind results. However, users can detect such cases since the size of the group is available in the L2B file output. Also, users are able to use the estimated standard error of each wind result to reject those that are estimated to be too noisy. The estimated HLOS standard errors are shown in Figure 15; as expected, the standard error is clearly larger in observations from the third group of measurements.

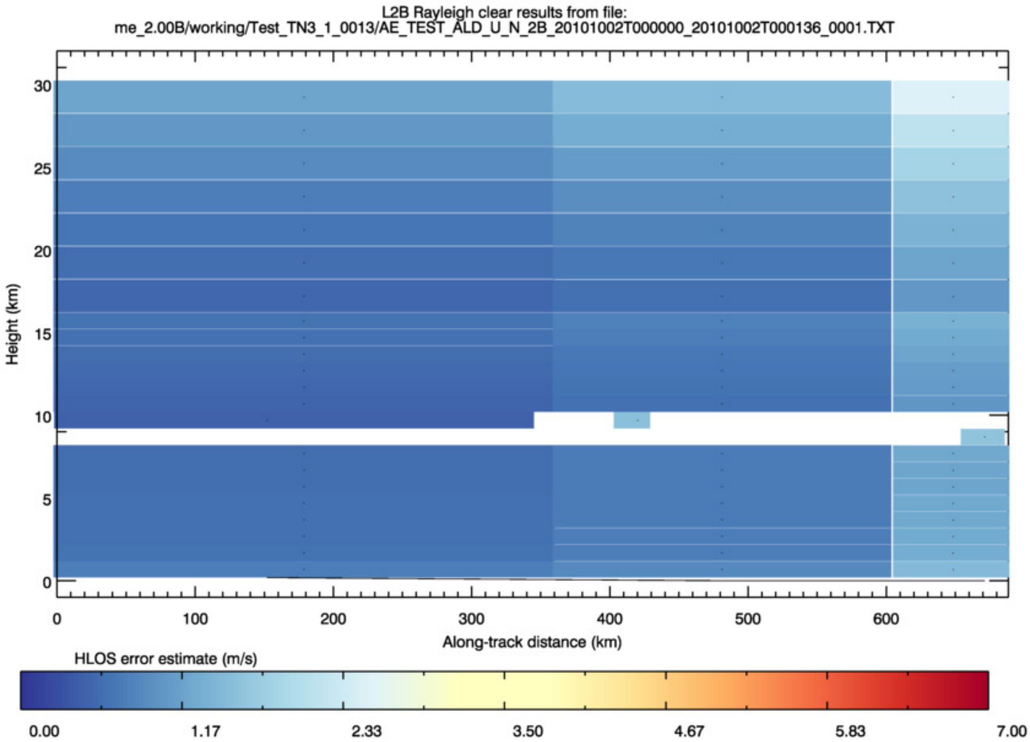


Figure 15. Estimated Rayleigh-clear HLOS standard error (m/s) as produced by the L2B processor for the test case of Figure 14.

4 Update of L2B/C file format

In the process of developing the grouping algorithm, KNMI suggested a rewrite of the L2B/C file format to better match what Aeolus products actually represent and to reduce the amount of missing data present in the products. This was accepted by the L2B team and has been implemented in the L2Bp v2.00. The new L2B/C file format is described in detail by a new IODD, see [RD 6]. Here we provide a brief overview.

The new file format is much simpler than the old one. It no longer consists of three nested loops over BRCs, Profiles/Classes and range-bins. With CM, since we no longer need to be restricted to the BRC definition, the product is no longer based on BRCs, but instead it is organised along individual wind retrievals. Also, without the artificial forcing of profiles of specific wind classes, the number of missing wind values can be greatly reduced. For example, a profile of Mie-cloudy results will rarely form a vertical profile covering all range-bins, due to the nature of the atmospheric attenuated backscatter from layers of cloud/aerosol, so profiles will have many missing values. Similarly, Rayleigh-cloudy profiles will by nature consist of mainly missing values.

The new file format does not enforce how the wind results are constructed, i.e. they could come from any imaginable combination of measurements allowing for great flexibility for future updates of the L2B processor. However, wind profile definitions are still possible. They are part of the new file format but are optional. Profiles may still be required if vertical wind shear assimilation is required (some way of tying together the wind results is necessary). Also, scientific researchers using L2B products will, no doubt, prefer vertical wind profiles. Note that ECMWF currently do not require vertical profiles of winds, but simply wind results with their associated geolocation information. Aeolus winds are assimilated as individual wind results, in the same manner as for aircraft winds.

In addition, the results of the grouping algorithm and AUX_MET data screening are stored in the new format. For traceability, each wind result and each profile has an identification code (identical to the Data Set Record index in the file). Finally, the connection of the wind results to the L1B data is made by a measurement map that holds an ID pointing to a wind result for each L1B measurement on each range-bin level.

5 Suggestions for further improvements in the Grouping Algorithm

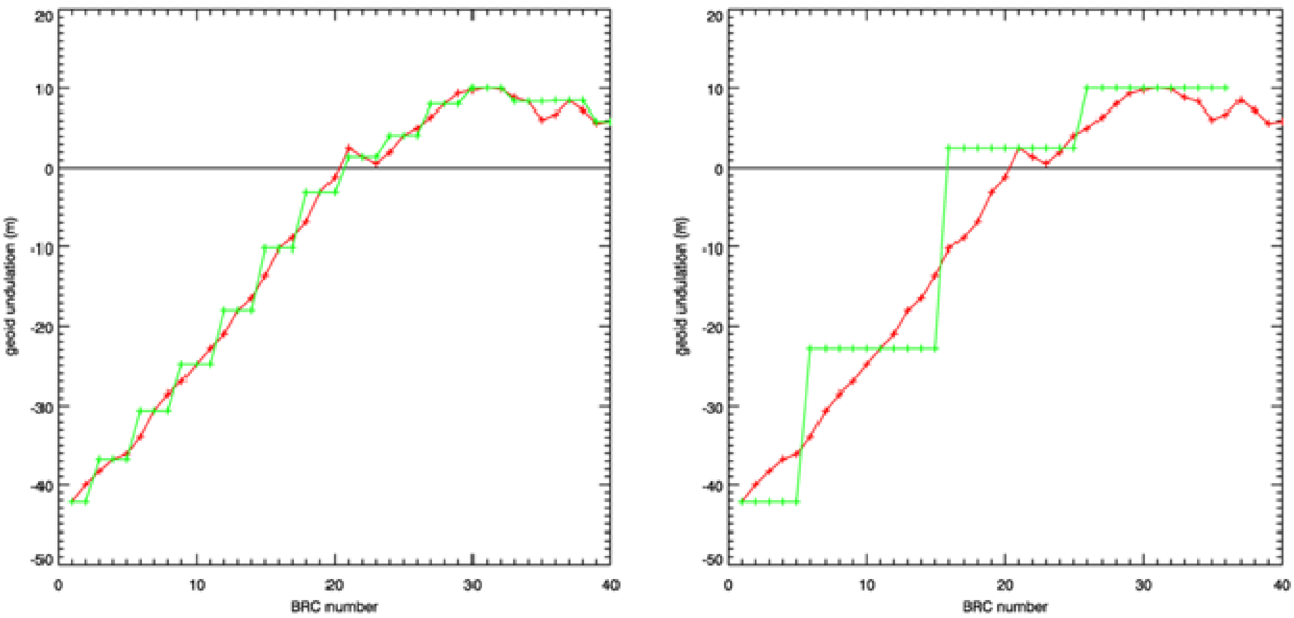
5.1 BRC constant variables

Some L1B variables have a constant value over a BRC, such as the geoid height and the zero wind correction (ZWC). At present, the L2B processor selects for the group the value assigned to a measurement in the centre of the group (a weighted average of measurement indices in the group, constant weighting in current L2Bp). If the group is particularly large, there is the potential for inappropriate values to be used for the wind result. A potential solution would be to implement a weighted average of the measurement level values across the group, rather than taking one measurement value from a weighted average of the indices.

Firstly we investigate the magnitude of this effect for geoid height. We present an arbitrary 40 BRC test case. Figure 16 shows the L1B reported geoid height in red (one value per BRC) and in green the central value that would be assigned for a group, for groups of a) 3 BRCs and b) 10 BRCs. For a given BRC number, the difference between the green a red lines indicates the size of the geoid height error (at the BRC level). Clearly, the error is larger at the edges of the group and when the geoid height is varying the most along the orbit track.

One can see that the geoid height error can be up to ~12 m for the 10 BRC case and ~4 m for the 3 BRC case at the edges of the group. However, when the geoid height is varying roughly linear across BRCs (as it does for the left half of the plot) then the chosen central value is effectively the mean value over the group, and so it is a sensible estimate. However when the geoid undulation is varying non-linearly (like for the right half of the plot), then the central value is not always a good choice, i.e. can be biased relative to the mean other values, by ~3 m for the 10 BRC case.

If there is vertical wind shear of say 10 [m/s]/km then the effective wind error due to the height assignment error of 3 m would in effect be a ~0.03 m/s wind error for the 10 BRC case, so in practice it is a small error (although it could induce vertical error correlations)



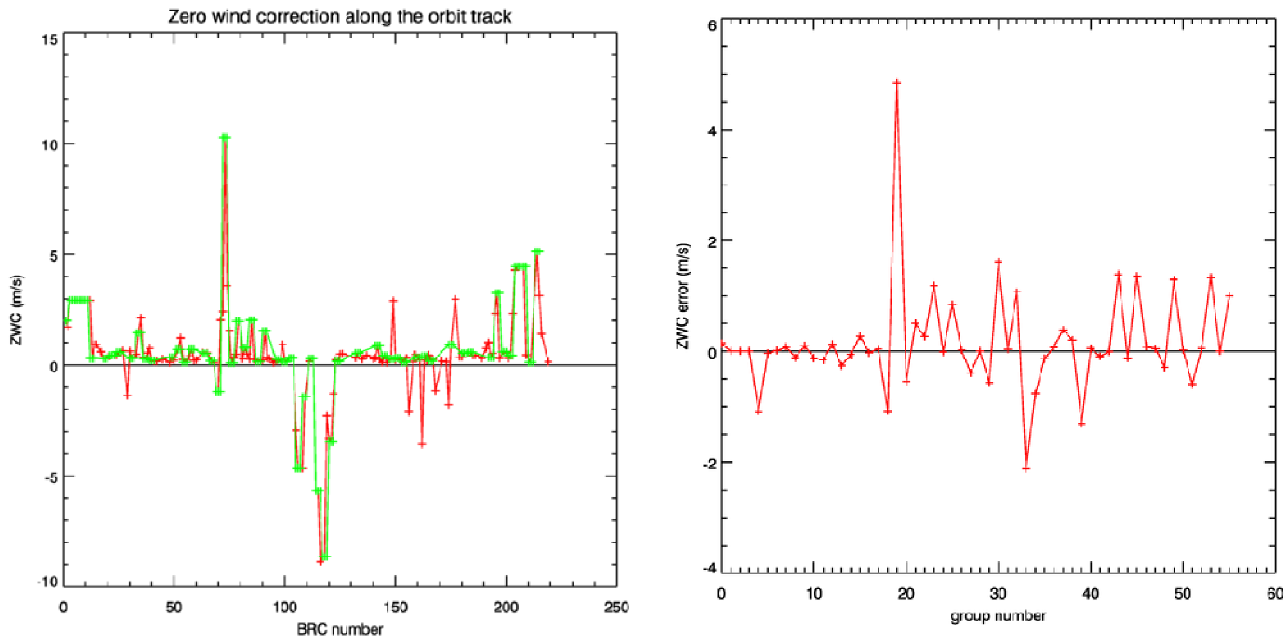
a) b)

Figure 16. The geoid height is plotted as a function of BRC number (from L1B data) in red. In green shows what effectively happens with the grouping algorithm, in that the central index of the observation value is applied. In a) the grouping is set to 3 BRCs, in b) the grouping is set to 10 BRCs.

A longer half-orbit of geoid heights was also looked at (not plotted here). It was found that the maximum error (between centre-of-group value and mean value over the group) for the 10 BRC grouping was ±4 m, so not much worse than our estimated 3 m estimate given for the shorter test case. Therefore, it appears that we do not need to worry about grouping sizes less than at least 10 BRCs as regards the geoid height. With groups of 30 BRCs, the errors start to exceed 10 m, which is too large (in effect producing wind errors of ~0.1 m/s in some cases). However, a 30 BRC group is unlikely to be needed in practice.

The same half-orbit test case (Mispointing_3_1_orb1_with_noise) was used to assess the ZWC variation along the orbit (in this case we set ZWC=*Mie_ground_corr_velocity* for each BRC from the L1B file). The test case uses CALIPSO backscatter and ECMWF fields of wind and temperature, and has all the E2S noise terms switched on (but no miss-pointing). The ground albedo is set to a constant high value, therefore ensuring relatively strong ground returns and hence potentially good ZWC values.

Figure 17 a) (similar to Figure 16) shows that the adjacent BRC variability of ZWC is much larger than the geoid height variation (at least for this test case). This variation is apparently due to noise. Therefore this poses a definite risk of error when assigning a group value of ZWC. Figure 17 b) shows the size of the error, which we define as the difference between the centre-of-group ZWC value and the mean-of-group value. With a relatively small group size of 3 BRCs, the errors can be fairly large on occasions (several m/s), due to the sharp variations in ZWC from BRC to BRC. This suggests it might be better to implement a weighted mean of the measurement ZWC values for observations created using groups larger than one BRC.



a) b)

Figure 17. a) The Mie ground correction velocity (which could be used as the ZWC in the L2Bp) for each BRC along the orbit. The red line is the value on each BRC (from the L1B file), and the green line is effect of choosing a centre-of-group value when group=3 BRCs (like with the current L2Bp algorithm) b) ZWC error (= the difference between mean of ZWC values in a group and the centre-of-group value) against group number, with group=3 BRCs.

We conclude that individual ZWC results based on a single BRC ground echoes are not suitable to be directly applied (they appear to be too noisy). It appears that some averaging, for example via a multi-BRC group at L2Bp level, or in the harmonic bias estimator (HBE) tool should be applied before trying to use these calibration corrections. HBE values should be smoothly changing along the orbit (by definition), so that 10 BRC groups should be fine with the current grouping algorithm, however we do not have HBE output to back this up.

5.2 Accumulation after grouping

Note that any suggestions of how to weight measurement-bins to produce wind observations, which follows after grouping and classification (e.g. using the estimated SNR values), are not an issue for the grouping algorithm per se, and so suggestions along those lines will not be listed here.

6 Conclusions

This technical note describes the grouping algorithm that has been developed for the CM L2B processor. Flexible horizontal resolution L2B winds (which are easily controlled) are now available with the CM L2B processor, meeting the requirements of NWP users of Aeolus data. It is also shown that the Grouping Algorithm functions as expected.

Some concern over how to apply the ZWC in the L2B processor was highlighted in this report, due to the fact that the Mie ground correction appears to be very noisy, and hence perhaps some averaging of ground returns are needed before applying such a correction.

Further improvements in how the measurement-bins are accumulated to construct wind observations within a group of measurements are possible, which could be implemented in a future L2B processor as part of work under a new contract.