AIRS/AMSU/HSB Version 7 Changes from Version 6

Edited by:
Heidar Th. Thrastarson[1]

Contributions by:
Eric Fetzer, Bjorn Lambrigtsen, Evan Manning,
Qing Yue, Baijun Tian
[1]Jet Propulsion Laboratory, California Institute of Technology
John Blaisdell, SAIC/GSFC, Joel Susskind, GSFC
Adam Milstein, MIT Lincoln Laboratory

29 July 2020
Document Version 1.0.0

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA

Submit Questions to:
https://airs.jpl.nasa.gov/data/support/ask-airs

Document Revision History

Date	Document Version	Changes
2 June 2020	0.3.2	Initial Release.
29 July 2020	1.0.0	Appendix added with field replacements. Links updated.

Table of Contents

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

Introduction and Overview

This document summarizes the main changes introduced in the AIRS/AMSU/HSB Version 7 (V7) release compared to the previous Version 6 (V6). This applies to Level 2 and Level 3 products from the AIRS retrieval algorithm. No significant changes have been made to the Level 1B processing for V7, and Level 1B products will continue to contain the V5 label. Level 1C products have been released under Version 6.7 and are described elsewhere.

Since the release of AIRS V6, significant improvements and modifications have been applied to the AIRS retrieval algorithm, especially the IR-only, which continues to provide high quality observations and retrievals after the complete loss of AMSU-A2 on September 24, 2016. Major algorithm changes between versions 6 and 7 are discussed in Section 2 and include:

- an improved Stochastic Cloud Clearing Neural Network (SCCNN) used as a first guess;
- removal of ambiguity in surface classification in the infrared-only (IR-only) retrieval algorithm;
- algorithm improvements that lead to better ozone, water vapor, and temperature products.

For a detailed description of the V7 algorithm, please refer to the V7 Algorithm Theoretical Basis Document:

V7_L2_ATBD.pdf

All AIRS documents can be obtained at the NASA Goddard Earth Science Data and Information Services Center (GES DISC) at:

https://disc.gsfc.nasa.gov/information/documents?title=AIRS%20Documentation

A more detailed characterization of the differences between V6 and V7 of various AIRS core data products can be found in the V7 testing and validation report:

V7_L2_Performance_Test_and_Validation_Report.pdf

This includes comparisons to other well validated data sources such as radiosondes, surface station, and satellite measurements. Highlights of Version 7 are as follows.

- Analysis of AIRS V6 water vapor products after its release revealed spurious bias in day-time products and large difference in the retrieval performance between day- and night-time cases. That bias has been removed in V7 as a result of removal of shortwave (2607~2657 cm^{-1}) channels used in the water vapor retrieval of both profiles and column integrals due to larger errors and higher uncertainties in the radiative transfer calculations over these channels.
- With the loss of AMSU-A2 and several AMSU-A1 channels, one of the goals with V7 has been to improve the quality of infrared-only algorithm, which continues to provide high quality global retrievals with reduced dependence on microwave observations. The current V7 infrared-only product uses a SCCNN first guess that utilizes stable microwave channels of AMSU-A1, while its physical retrieval step only uses infrared channels and does not use microwave information. Plans are in place to transition to a pure infrared-only SCCNN first guess if there is further loss of microwave channels in the future.
- The ambiguity in differentiating frozen and non-frozen surface classification in the IR-only algorithm has been removed by using ice concentration over ocean and water equivalent accumulated snow depth over land obtained from the ancillary Global Forecasting System (GFS) files. We note that until the loss of AMSU-A2 in 2016, this information was derived from the microwave observations.
- Comparisons of temperature and water vapor vertical profiles with radiosonde and reanalysis data show an improved performance over all in V7.
- For ozone, V7 includes more channels sensitive to O_3 in the upper troposphere and lower stratosphere in the 9.6μm O_3 band and uses additional vertical basis functions to better represent the profile. The ozone retrieval also benefits from improved polar stratospheric temperatures in the V7 SCCNN and better quality control of the retrievals. Comparisons with OMPS and OMI both indicate an improved column ozone retrieval in V7.

Several issues in the V7 retrievals have been identified. These issues, which affect only a limited set of particular regions or conditions, have no discernible impact on global statistics. We note that there are larger errors in V7 SCCNN surface skin temperature over the Antarctic sea ice, the challenges of near surface profile retrievals over snow- and ice-covered surfaces, and ambiguity in cloud retrievals when effective cloud fraction is less than 5%, all of which are highly regional in nature.

The version numbers that appear in the AIRS Product Files are slightly different, depending upon the product due to a staged delivery of processing code to the GES DISC. They are:

- **Level 1B AMSU-A and HSB Products**: v5.0.0.0
- **Level 1B AIRS Products**: v5.0.x.0
- **Level 1B Calibration Subset Product**: v5.0.x.0 is still available. It will be superseded at some point in the future.
- **Level 1C:** v6.7.2.0
- **Level 2 Products**:
    - **Level 2 IR-Only**, **AIRS+AMSU** and **AIRS+AMSU+HSB**, standard, support, and cloud-cleared radiance products: v7.0.2.0
    - **Level 2 CO_2**; V5.4.11.0
- **Level 3 Products**:
    - **Level 3 Level 2 IR**-Only, **AIRS+AMSU** and **AIRS+AMSU+HSB,** standard, and support products: v7.0.3.0
    - **Level 3 CO_2**; V5.9.14.0

2 Summary of Algorithm Improvements in V7

Based on the evaluation of the performance of V6, numerous scientific modifications have been made to the retrieval algorithm for V7 to improve the retrieval quality. The V7 retrieval flow is described in

V7_Retrieval_Flow.pdf

This short document includes a step-by-step comparison of the V6 and V7 retrieval flow. The improvements in the retrieval algorithm from V6 to V7 are also covered in the V7 testing and validation report:

V7_L2_Performance_Test_and_Validation_Report.pdf

The summary of algorithm improvements contained therein is largely repeated here for convenience.

2.1 Changes to Stochastic Cloud Clearing Neural Network (SCCNN)

The SCCNN algorithm is a statistical method for retrieval of temperature and water vapor profiles as well as surface parameters using radiance observations from AIRS and AMSU. The algorithm consists of two key parts: SCC (Stochastic Cloud Clearing) estimates the cloud-cleared infrared spectrum using series of linear and nonlinear operations on AIRS/AMSU radiances from a 3×3 “golf ball” footprint. The NN (neural network) estimates temperature and water vapor profile from projected principal components of the cloud-cleared spectrum. SCCNN surface skin temperature, temperature and water vapor profiles are used as first guess in V6 and V7 AIRS retrievals and are also provided in the AIRS Level 2 Support product. The NN algorithm is a regression, trained once, then applied over the entire mission. Its products are generated in near-real-time. Because it is trained once and for all, the NN first guess is expected to behave similarly throughout the whole mission, while model-based first guesses can show significant change in bias structure over mission duration due to model changes.

A new SCCNN algorithm version was developed and delivered as the AIRS V7 IR+MW and IR-only algorithm first guesses. The new SCCNN algorithm has been trained on newer, more diverse training sets. The representation of the atmosphere used internally by SCCNN now uses 91 vertical levels rather than the previous version’s 60 levels, matching the more recent ECMWF fields it is (mostly) trained with. A series of versions with incremental changes during the V7 development are summarized as the following:

- New coefficients trained on 2013-era ECMWF (v6.46)
    - Addressed previously noted issues from prior V7 candidates (V6.28) related to training set interpolation artifacts, discontinuities in sea ice flagging, and training set endpoints that truncated a portion of the summer.
- Increased training set comprehensiveness to capture a wide range of variability (v6.50)
    - Now trained using data from 2013, 2010, and 2005.
- Improved principal component representation for SCC algorithm, which improved cloud clearing results in IR-only algorithm (v6.51).
- Improved SCCNN for V7 AIRS IR-only algorithm by including AMSU-A1 (v6.54/v6.55):
    - 6.5.5 SCCNN, for AIRS+AMSU-A1-only, uses AMSU channels 3, 8-13, 15. (V6 SCCNN, in comparison, used channels 1-3, 6, 8-15.), which provides an option for addressing the degradation and failure of AMSU+A2.
    - Removed flagged invalid channels from IR channel list used in SCCNN and replaced with other, valid channels that sound near the PBL.
- Updated training set screening in poles (v6.54/v6.55) which causes modest polar improvements versus v6.51 overall.

2.2 Changes to Surface Classification in the IR-only Algorithm

In the AIRS-Only system, V6 uses solely the criteria of whether surface temperature from SCCNN is above or below the freezing point to determine whether the surface is frozen or not, which is ambiguous near the freezing point and causes reduced retrieval performance (Yue et al. 2017[1]). The V7 IR-only algorithm determines the surface classification using the ice concentration over ocean and water equivalent accumulated snow depth over land from the Global Forecasting System (GFS).

2.3 Changes to the Ozone Retrieval Algorithm

A new climatological first guess was developed which separates ozone hole cases from other cases to yield two distinct profile shapes over Antarctica during

[1] Yue, Q., and B. Lambrigtsen, 2017, AIRS V6 Test Report Supplement: Performance of AIRS+AMSU vs. AIRS-only Retrievals, available at: https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/AIRS/3.3_ScienceDataProductDocumentation/3.3.5_ProductQuality/V6_Test_Report_Supplement_Performance_of_AIRS+AMSU_vs_AIRS-Only_Retrievals.pdf.

spring. On a profile-by-profile basis, the 50hPa temperature given by the SCCNN is used to select the profile shapes to use as first guess in the retrieval process. The final retrieval shows improved ozone profiles.

Outside the ozone hole the ozone first guess profile is adjusted to the tropopause given by the SCCNN. A shape-preserving transformation is made to align the climatological tropopause to the case-by-case tropopause in the ozone first guess, with no adjustment applied above 1hPa. The ozone first guess is now included in the V7 L2 product.

More trapezoids are added to allow better representation of the vertical structure where the information is available. The damping was decreased everywhere and was made even less during summer mid-latitude and polar months, recognizing that very large excursions from the first guess are normal in propagating waves during those months.

Many of the strongest ozone lines were added to the retrieval for V7 except in the relatively few very cold cases where there is little contrast between the stratosphere and the surface, channels between 1031 cm^{-1} and 1061 cm^{-1} are not used. In all other spatial regions these channels increase the degrees of freedom over V6 and improve the total ozone retrieval.

Two additional spectral hinge points in the 9.6 μm region were added to the surface retrieval (along with channels in this region), and their emissivity values are also updated in the ozone retrieval step.

The surface emissivity noise covariance contribution was removed from ozone retrieval, since emissivity is now part of the solution space. The water noise covariance contribution was also removed because it was found that this reduced the overdamping present in all previous versions.

Quality control specific to ozone was added. The ozone retrieval is rejected under the following conditions: the attempted change on the first iteration is too large indicating the temperature or water profiles may be in error, radiative closure is not achieved in the ozone channels, the dust test indicates the presence of significant dust, and/or the change in emissivity from the first guess is very different in the ozone spectral region and the adjacent spectral regions.

These algorithm improvements lead to an increase of information content and vertical resolution of the ozone product. Together with improved ozone quality control, these changes enabled a much better match to OMI and OMPS data (see **V7_L2_Performance_Test_and_Validation_Report.pdf**).

2.4 Changes to the Water Vapor Retrieval Algorithm

In V7, the shortwave channels in the range of 2607~2657 cm^{-1} are removed from the water retrieval due to larger errors and higher uncertainties in the radiative transfer calculations over these channels, which eliminated the spurious day-night bias reported in V6 water products (see **V7_L2_Performance_Test_and_Validation_Report.pdf**).

Several other changes were also made:

- The number of humidity trapezoids was increased to better represent the boundary layer humidity and the trapezoid boundaries were made a subset of the temperature trapezoid boundaries to increase the stability of the retrieval.
- The damping in the water retrieval was increased to reduce the degradation of near surface water vapor profiles relative to SCCNN results.
- Channel changes: The strongest water lines were removed from the retrieval as it is shown their utilization had a negative effect. Water continuum channels were added to the water retrieval following the suggestion of the science team.
- In the water vapor first guess, we now taper the SCCNN profile to climatology at four levels above the tropopause. This has removed spurious structure found in the stratospheric water profile present in V6.

2.5 Changes to the Temperature and Surface Retrieval Algorithm

- The temperature retrieval trapezoids were increased in the boundary layer and were made more uniform vertically, which led to a better representation of the boundary layer and of the tropopause.
- Longwave channels with cloud clearing corrections larger than 5K for a given profile are now excluded from the temperature retrieval for that profile.
- In V6, the surface temperature was modified in the temperature retrieval step. This feature was removed for V7.
- Seven additional 4μm channels were added to the temperature retrieval to increase information at the top of the atmosphere, which also affects the ozone retrieval.
- The emissivity channel list was modified to ensure information in the ozone band is used to adjust the surface emissivity there.

2.6 Changes to Temperature and Water Vapor Quality Control (QC)

The following modifications have been made:

- The error estimate used as the decision point for atmospheric temperature and water profile QCs is moved from 6 layers above the surface in V6 to 2 layers above the surface over frozen and land surfaces in V7. This effectively uses the entire profile error estimate information and allows the algorithm to make finer distinctions of quality over land and frozen areas near the surface.
- The numerical threshold points for marking profile levels with QC=2 is tightened over land from the mid to lower atmosphere, and for frozen cases in the middle atmosphere, while the numerical threshold for frozen cases near the surface is slightly loosened which increases the yield over frozen surfaces.
- One relative humidity quality check was modified. In V6, cases with unrealistic relative humidity and low cloud fractions are identified and the profiles are marked as QC=2 up to 100hPa. In V7 this test is revised to address the uncertainty associated with the cloud top pressure in a less restrictive way by moving PGood and PBest only a few levels instead of to 100hPa.

3 V7 L2 Product Improvements, Changes, and Performance Limitations

Details on major improvements, changes, and performance limitations of Level 2 products are covered in

V7_L2_Performance_Test_and_Validation_Report.pdf

3.1 Major Improvements in the V7 Retrieval Products

The improvements covered in the testing and validation report include the following:

- Better agreements between IR-only surface classification and NSIDC data
- Removal of spurious day-time bias and day-night performance difference on water vapor retrieval
- New Stochastic Cloud Clear Neural Network (SCCNN) for V7 AIRS IR-only algorithm by including AMSU-A1
- Smaller biases and Root Mean Square Errors (RMSEs) in PBL over non-frozen ocean. Over non-frozen land, the bias in the V7 water vapor profiles is smaller than for V6, which is mainly due to an improved first guess from SCCNN.
- For ozone products, Version 7 shows significant improvements in vertical resolution and agreement with ozonesonde in the lower stratosphere. The bias in total column ozone is greatly reduced in V7 than in V6 by comparison with OMI and OMPS products.

3.2 Performance Limitations of Challenging Scenes in V7

During the initial testing of the V7 product, despite the overall improvements of the V7 retrieval products, several performance limitations have been identified. These issues are related to highly challenging scenes such as surface temperature over Antarctic sea ice, near-surface temperature and water vapor profiles over snow and/or ice-covered land, and cloud properties in cases with small effective cloud fractions (ECF < 0.05), all of which are highly regional in nature. Analyses of these scenes are presented in detail in the V7 testing and validation report. More thorough studies are needed to examine how these known limitations impact science investigations using AIRS data in these conditions. It is also noted that the values of degree of freedom (DOF) and verticality of the water vapor and temperature profiles reported in the V7 Level 2 product are smaller than those in V6. This is related to how the current algorithm (both V6 and V7) calculates the temperature and water vapor averaging kernel matrices and error estimate and users should be cautious using these variables interpreting the sensitivity and information content of the AIRS spectra.

4 V7 L3 Product Changes from V6

There are five major new improvements or changes for V7 L3 products in comparison to V6 L3 products. Further descriptions are included in

V7_L3_Product_User_Guide.pdf

The monthly averaging method has been changed, some ancillary fields removed, the 8-day and quantized products removed, and total column CO and CH4 products and surface O_3, CO and CH_4 products have been removed.

4.1 Change in the Monthly Averaging Method

The AIRS V6 L3 monthly product is generated by the “Averaged By Observation” (ABO) approach. However, the AIRS V7 L3 monthly product is generated by the “Averaged By Day” (ABD) approach. The ABO approach is a “weighted” average of daily means, where the counts of successful retrievals in a 1 × 1 degree grid cell for each day serve as weights. The days with more valid counts have more weights in the monthly average. The day-to-day variations of valid counts arise from the orbit shift and retrieval algorithm. Since the AIRS retrieval fails in conditions with cloud cover more than 70%, the counts will consistently weigh in favor of less-cloudy conditions and cause sampling bias in the V6 L3 standard monthly product. The sampling bias may become more significant and consistent in multi-year climatological estimates. To alleviate this skewed sampling, the ABD approach is to derive the monthly averages from the daily means, without regard to the counts of observations.

4.2 Removal of Less Used Ancillary Fields

For the AIRS V6 L3 product, there are also up to 5 ancillary fields to complement each retrieval mean (without any suffix): standard deviation (with a suffix of _sdev), minimum (with a suffix of _min), maximum (with a suffix of _max), input count (with a suffix of _ct), and error (with a suffix of _err). For the AIRS V7 L3 product, there are only 2 ancillary fields to complement each retrieval mean (without any suffix): input count (with a suffix of _ct) and standard deviation (with a suffix of _sdev). This has significantly reduced the file size of the AIRS V7 L3 products.

4.3 Removal of Less Used 8-Day Product

For the AIRS V6 L3 product, there are three temporal resolutions: daily, 8-day (half of the 16-day Aqua orbit repeat cycle) and monthly (calendar). In contrast, there are only two temporal resolutions: daily and monthly (calendar) for the AIRS V7 L3 product. The less-used 8-day L3 product has been removed from the V7.

4.4 Removal of Less Used L3 Quant Product

For the AIRS V6 L3 product, there is a 5°x5° L3 quantized product (L3Q) that summarizes mean, standard deviation and number of observations on a 5x5 degree grid per pentad (5 days) and calendar month for up to 100 clusters at 10 pressure levels. Day and night data is mixed together. For the AIRS V7 L3 product, this 5°x5° L3 quantized product (L3Q) has been removed due to its infrequent usage. This has significantly reduced the file size of the AIRS V7 L3 products.

4.5 Removal of Total Column CO and CH_4 and Surface O_3, CO and CH_4 Products

The peak sensitivity of the AIRS retrieval to carbon monoxide (CO) occurs at 500 hPa and the peak sensitivity of the AIRS retrieval to methane (CH_4) occurs at 300 hPa. AIRS has no skill in retrieving the CO and CH_4 surface amounts, and the information we report on the total column CO and CH_4 comes almost entirely from the climatology used in the first guess. For these reasons we have removed the total column and surface CO and CH_4 reported in the V6 L3 product from the V7 L3 product. Only the mid-tropospheric CO and CH_4 mixing ratios are kept in the V7 L3 product. Similarly, AIRS has no skill in retrieving the ozone (O_3) surface amounts and we have removed the surface O_3 reported in the V6 L3 product from the V7 L3 product.

Appendix: Field Replacements

For some variables, the sizes of arrays have changed from Version 6 to Version 7. In these cases, a string has been added to the variable name, indicating the new dimension size, so that user's Version 6 analysis software will not inadvertently use the new, differently dimensioned output. The number of trapezoidal layers used in retrievals has been increased in Version 7, from 23 to 30 for temperature retrievals, from 15 to 21 for water vapor, and from 9 to 20 for ozone retrievals. The table below lists the affected fields and their names in both versions.

Version 6 Field Name	Version 7 Field Name
Temp_trapezoid_layers	**Temp_trapezoid_layers_30func**
Temp_eff_press	**Temp_eff_press_30func**
Temp_ave_kern	**Temp_ave_kern_30func**
Temp_verticality	**Temp_verticality_30func**
H2O_trapezoid_layers	**H2O_trapezoid_layers_21func**
H2O_eff_press	**H2O_eff_press_21func**
H2O_ave_kern	**H2O_ave_kern_21func**
H2O_verticality	**H2O_verticality_21func**
H2O_VMR_eff	**H2O_VMR_eff_21func**
H2O_VMR_eff_QC	**H2O_VMR_eff_21func _QC**
H2O_VMR_eff_err	**H2O_VMR_eff_21func _err**
O3_trapezoid_layers	**O3_trapezoid_layers_20func**
O3_eff_press	**O3_eff_press_20func**
O3_ave_kern	**O3_ave_kern_20func**
O3_verticality	**O3_verticality_20func**
O3_VMR_eff	**O3_VMR_eff_20func**
O3_VMR_eff_QC	**O3_VMR_eff_20func_QC**
O3_VMR_eff_err	**O3_VMR_eff_20func_err**