ALGORITHM THEORETICAL BASIS DOCUMENT AIRS AMMONIA RETRIEVAL

Prepared by

Juying Warner, University of Maryland (UMD)

Contributions by:

Zigang Wei, UMD

Lena Iredell, NASA GES DISC

Version 3

February 2019
1. INTRODUCTION .................................................................................................................. 3

1.1. Atmospheric Infrared Sounder (AIRS) Retrieval Algorithm ........................................ 3

1.2. Abbreviations and Acronyms .......................................................................................... 5

2. AIRS L2 DATA PRODUCTS used in \( \text{nh}_3 \) retrievals .................................................. 7

2.1. Cloud-Cleared Radiances ............................................................................................... 7

2.2. AIRS L2 Support Products ............................................................................................ 8

2.3. AIRS Errors and Noises Used ....................................................................................... 8

3. AIRS \( \text{NH}_3 \) RETRIEVAL METHOD .............................................................................. 9

3.1. Steps Used in the Retrievals in Version 3 ..................................................................... 9

3.2. Formulation ................................................................................................................... 10

3.3. Channels Used in the Retrieval Steps .......................................................................... 11

3.4. A Priori ......................................................................................................................... 12

4. QUALITY ASSURANCES AND ERROR ESTIMATES USED IN Version-3. 14

4.1. Quality Assurances ..................................................................................................... 14

4.2. Error Estimates ............................................................................................................ 15

5. DATA SETS .................................................................................................................... 16

5.1. Monthly Data Sets ....................................................................................................... 16

5.2. Daily Data Sets ........................................................................................................... 17

6. REFERENCES .................................................................................................................. 17
1 INTRODUCTION

1.1 Atmospheric Infrared Sounder (AIRS) Team Retrieval Algorithm

The Atmospheric Infrared Sounder (AIRS) is a cross-track high spectral resolution infrared sounder onboard NASA’s Aqua satellite that channel collects radiance data with a 13.5 km spatial resolution in the horizontal at nadir. It has 2378 infrared channels from which the AIRS Science Team generates highly accurate Quality Controlled temperature and humidity profiles, trace gas profiles, and surface parameters. AIRS is an imaging hyperspectral sounder, which covers 80% of the globe twice per day, during the ascending (day) and the descending (night) overpasses. However, the images are in scan coordinates, with considerable overlap at high latitudes and gaps near the equator.

The AIRS Science Team retrieval algorithm generates surface and atmospheric parameters using AIRS radiance observations taken within the single 3x3 array of AIRS Fields of View (FOV’s) which lies within an AMSU A footprint, called the AIRS Field of Regard (FOR). Retrievals of most geophysical parameters are performed on an AIRS FOR basis, which has a spatial resolution of 40 km at nadir. Cloud products are retrieved on an AIRS FOV basis. The AIRS team retrieval process consists of four major components. The first component uses observed AIRS channel radiances within the AIRS FOR to generate the Neural Net initial guess (Milstein 2015) used to start the retrieval process. The second component is the generation of clear column radiances in the FOR for all channels, which would have been observed if there were no clouds contained within the AIRS FOR. The third component is a multi-step physically based retrieval procedure, starting with the initial guess, that finds geophysical parameters which best match both the observed AMSU radiances and AIRS clear column radiances for those
channels used in a given step of the retrieval process. Retrieval steps are done sequentially and determine the following geophysical parameters: surface skin temperature, surface spectral emissivity, surface bidirectional reflectance; atmospheric temperature profile, atmospheric moisture profile, and also other geophysical parameters in the AIRS FOR, as well as the fractional cloud cover in each AIRS FOV. These steps are done sequentially so as to make the retrieval process as linear as possible in each step and allow for use of a set of channels in each retrieval step whose radiances are most sensitive to what is being solved for in that step and parameters previously solved for, while being relatively insensitive to geophysical parameters not yet solved for. Additionally, unlike most other retrieval methodologies, there is no explicit weight given to either an a-priori state or the initial guess.

The basic cloud clearing and retrieval methodologies used in the AIRS Science Team retrieval algorithm, including the meaning and derivation of Jacobians, the channel noise covariance matrix, and the use of constraints including the background term, are essentially identical to the AIRS Science Team pre-launch algorithm described in Susskind et al. (2003). AIRS Radiative Transfer Algorithm (RTA) SARTA was described by Strow et al. (2003). The current AIRS Version-6 retrieval algorithm contains many further improvements in retrieval methodology.

This document describes the details of the theoretical basis of the AIRS NH$_3$ products algorithm. The AIRS Level 2 products algorithm, not including AIRS NH$_3$, is documented and distributed as the AIRS Level 2 Product Generation Executive (PGE), Version 6.0, at the Goddard Space Flight Center (GSFC) GES DISC.
1.2 Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRS</td>
<td>Atmospheric Infrared Sounder</td>
</tr>
<tr>
<td>AK</td>
<td>Averaging Kernels</td>
</tr>
<tr>
<td>AO</td>
<td>AIRS Only</td>
</tr>
<tr>
<td>SARTA</td>
<td>Stand-alone AIRS Radiative Transfer Algorithm (RTA)</td>
</tr>
<tr>
<td>AMSU</td>
<td>Advanced Microwave Sounding Unit</td>
</tr>
<tr>
<td>ATBD</td>
<td>Algorithm Theoretical Basis Document</td>
</tr>
<tr>
<td>CRDS</td>
<td>Cavity Ring Down Spectrometer</td>
</tr>
<tr>
<td>CCR</td>
<td>Cloud-Cleared Radiances</td>
</tr>
<tr>
<td>CIMS</td>
<td>Chemical Ionization Mass Spectrometry</td>
</tr>
<tr>
<td>COLR</td>
<td>Clear Sky Outgoing Longwave Radiation</td>
</tr>
<tr>
<td>DBTI</td>
<td>Difference of Brightness Temperature Index</td>
</tr>
<tr>
<td>DISCOVER-AQ</td>
<td>DISCOVER-AQ Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality</td>
</tr>
<tr>
<td>DOFS</td>
<td>Degrees of Freedom for Signal</td>
</tr>
<tr>
<td>EOS</td>
<td>Earth Observing System</td>
</tr>
<tr>
<td>FOR</td>
<td>Field of Regard</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of View</td>
</tr>
<tr>
<td>FM</td>
<td>Forward Model</td>
</tr>
<tr>
<td>GEOS</td>
<td>Goddard Earth Observing System</td>
</tr>
<tr>
<td>GSFC</td>
<td>Goddard Space Flight Center</td>
</tr>
<tr>
<td>HITRAN</td>
<td>High Resolution Transmission Molecular Absorption Database</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
</tbody>
</table>
kPa  kilopascal \((10^3\) pascal, equivalent to 10 bar) 

L1  (Level-1) Instrument data at full resolution, time-referenced, and annotated with ancilliary information, including radiomeric and geometric calibration coefficients and georeferencing parameters that have been processed to sensor units.

L2  (Level 2) Derived geophysical variables at the same resolution and location as Level 1 source data.

L3  (Level 3) Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.

MERRA  Modern-Era Retrospective analysis for Research and Applications

MODIS  Moderate Resolution Imaging Spectroradiometer

NASA  National Aeronautics and Space Administration

NEDT  Noise Equivalent Differential Temperature

OE  Optimal Estimation

OLR  Outgoing Longwave Radiation

QC  Quality Control

SCC/NN  Stochastic Cloud Clearing/Neural Network

TC  Thermal Contrast

TES  Tropospheric Emissions Spectrometer

VMR  Volume Mixing Ratios
2 AIRS L2 DATA PRODUCTS USED IN NH₃ RETRIEVALS

AIRS granules are formally defined as the smallest aggregation of data that is independently managed (i.e., described, inventoried, retrievable). An AIRS granule has been set as 6 minutes of data, corresponding to exactly 45 scanlines of AMSU data or 135 scanlines of AIRS data. The orbit repeat-pattern of the EOS Aqua is 16 days, i.e. the spatial coverage of the N-th granule is repeated (almost) exactly 16 days later. There are two types of AIRS L2 products: Standard Products and Support Products. We use AIRS L2 Support Products (i.e., meteorological and trace gas profiles, and surface parameters) as input to the NH₃ retrieval codes. The Support Product includes higher vertical resolution profiles of the quantities found in the Standard Product, but at all locations as Standard Products.

2.1 Cloud-Cleared Radiances

Cloud-Cleared Radiances are produced along with the AIRS L2 Product, as they are the radiances used to retrieve the Standard Product. They are a separate output file, but like the Standard Product, are generated at all locations. For detailed algorithm description, see Suskind et al. (2003), and refer to AIRS L2 ATBD at NASA GES DISC. (https://disc.gsfc.nasa.gov/datasets/AIRX2SUP_V006/summary?keywords=AIRS%20L2%20Support)

<table>
<thead>
<tr>
<th>Product Short Name</th>
<th>Product Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRICCF</td>
<td>CC</td>
<td>Level-2 cloud cleared radiance product created using AIRS IR, AMSU without-HSB</td>
</tr>
</tbody>
</table>
2.2 **AIRS L2 Support Products**

The Support Product profiles contain 100 levels between 1100 and .016 hPa; this higher resolution simplifies the generation of radiances using forward models, though the vertical information content is no greater than in the Standard Product profiles. AIRS produces a number of meteorological products, i.e., temperature and water vapor profiles, clouds and surface quantities, that were developed and tested. Also, among these are trace constituent profiles of CO and CH$_4$, Outgoing Longwave Radiation (OLR) and Clear Sky Outgoing Radiation (COLR), and CO$_2$ product. The profiles and surface parameters from the Support Products are used as input in the AIRS forward model, SARTA (Strow et al., 2003), in the NH$_3$ retrieval codes.

<table>
<thead>
<tr>
<th>Product Short Name</th>
<th>Product Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRX2SUP</td>
<td>RetSup</td>
<td>Level-2 retrieval support product created using AIRS IR, AMSU without-HSB</td>
</tr>
</tbody>
</table>

2.3 **AIRS Errors and Noises Used**

In addition to the instrument noise at individual L1 radiance channels (NEDT), we need to consider the noise introduced in the cloud-clearing process in AIRS NH$_3$ retrievals. We also need to consider the forward model uncertainties in SARTA and in the input profiles, i.e., AIRS L2 profiles. The CCR and L2 profile noises are produced in the AIRS L2 retrieval processes.

Finally, version-6 has an “AIRS Only” (AO) processing capability which utilizes only AIRS observations and produces results only slightly degraded from those obtained utilizing both AIRS and AMSU observations. This “AIRS Only” capability is an important backup processing mode because noise performance on some channels of AMSU-A is continuing to
degrade, and at the end of August 2016, the version-6 algorithm switched completely to AIRS Only. The AIRS NH₃ data record stops at this point, and more detailed study is required at a later time.

3 AIRS NH₃ RETRIEVAL METHOD

3.1 Retrieval Steps Used in AIRS NH₃ Version 3

The retrieval steps for AIRS NH₃ are illustrated in Figure 1, which summarizes the input and the output properties and the main retrieval module. The meteorological parameters, such as pressure, temperature, the profiles of constituents other than NH₃, and the surface parameters, are retrieved in the earlier steps of the AIRS operational retrieval system and used as input to the forward model (FM), SARTA (Strow et al., 2003), which is an integral part of the retrieval system, (see subsections 2.2 and 3.2). We use the Optimal Estimation (OE) method in AIRS NH₃ retrievals by following the formulations given by Rodgers (2000); see details of the minimization routines in
subsection 3.2. Cloud-cleared radiances from AIRS L2 products are used instead of AIRS L1 data, see subsection 2.1. A priori information is critical in the OE retrieval system, see detailed description in subsection 3.4

### 3.2 Formulation

This section discusses the OE method used in the AIRS NH$_3$ retrievals. Given a model of the instrument’s signals, the FM for the CO profile retrieval problem can be written as:

\[ y = f(x, b) + n^" \]

where \( y \) is the vector of measured radiances, \( x \) is the state vector (variables to be retrieved from the measurements), \( b \) represents all other parameters used by the forward model, \( f(x, b) \) is the forward model function, and \( n^" \) is the instrument noise. AIRS L2 profiles and Surface parameters that are inputted to the forward model are represented as \( b \). The CCR is represented by \( y \); NH$_3$ is in \( x \); and \( f(x, b) \) follows SARTA. The retrieval process inverts Eq. (1) to find \( x \) for a given \( y \). As in the case of atmospheric remote soundings, the inverse problem is usually ill posed, because the number of independent measurements is less than the number of variables to be retrieved. The OE method introduces the a priori information as an additional constraint. The solution then can be understood as the combination of the present measurements and the prior knowledge.

For the variables that obey a Gaussian distribution, this inverse problem is equivalent to the maximum likelihood solution, and by using a Newtonian iteration, the solution to Eq. (1) can be written as:
\[ x_{n+1} = x_a + C_a K_n^T (K_n C_a K_n^T + C_e)^{-1} [y - y_n - K_n (x_a - x_n)] \]  

where \( n \) is the order of iteration and \( C_e \) is the measurement error covariance matrix. \( K_n \) is the Jacobian matrix for iteration \( n \), \( x_a \) is the mean of the a priori distribution, and \( C_a \) is the a priori covariance matrix for \( x_a \). We used the a priori matrix developed from the GEOS-Chem model in this study.

It is important to apply the averaging kernel information in the sensitivity analyses using the OE method. As defined by the retrieval formulations, the averaging kernels are computed using the following:

\[ A = C_a K^T (K C_a K^T + C_e)^{-1} K \]  

and,

\[ x_0 = A x + (I - A) x_a, \]  

where \( I \) represents the identity matrix and \( x \) is the true state. Equation (4) states that in the absence of other error sources the retrieved state is a weighted mean of the true state and the a priori state, with the weight \( A \) for the true state and \( I - A \) for the a priori. This shows the importance of averaging kernels as diagnostics of the retrieval. The closer the matrix \( A \) is to the identity matrix the better the retrieved state resembles the true state. Equation (4) also shows that unless the matrix \( A \) is an identity matrix, each layer of the retrieved state is an averaged contribution of multiple layers in the true state and the a priori state.
3.3 **Channels Used in the Retrieval Steps**

Figure 2 left panel shows a typical AIRS cloud-cleared brightness temperature spectrum in black lines, where the over-plotted red dots are those channels used to determine NH$_3$ a priori levels and to retrieve NH$_3$ VMRs. Figure 2 right panel shows AIRS channel sensitivities for all species in the NH$_3$ spectral region. The red channels show NH$_3$ sensitivities and the black dots on the red channels are what we used in the NH$_3$ retrievals. Channels shown in other colors: green, blue, brown, and purple, are the sensitivities for carbon dioxide, water vapor, ozone, and nitric acid, respectively.

![AIRS Sensitivities, Lat=15.0°](image)

Figure 2 left panel. a typical AIRS cloud-cleared brightness temperature spectrum in black lines, where the over-plotted red dots are the channels used for NH$_3$. Figure 2 right panel. AIRS channel sensitivities for all species in the NH$_3$ spectral region.
3.4 *A Priori*

The optimal estimation method requires an a priori mean profile and a corresponding error covariance matrix that represent the current knowledge of the geophysical property, i.e., NH$_3$, prior to the retrieval. Due to the high spatial variability and short lifetime of NH$_3$, a simple fixed a priori for all emission scenarios is not appropriate. We developed a global mean, multi-year averaged (2003–2012), three-level a priori from GEOS-Chem model (v9-02) simulations for high, moderate, and low emission levels. We used GEOS-5 MERRA datasets from the NASA Global Modeling and Assimilation (Rienecker et al., 2011) to drive the meteorological fields in the GEOS-Chem simulations. Figure 3 shows the a priori mean profiles (solid curve with squares)

![Figure 3. The a priori profiles and the square root of the diagonal terms of the error covariance matrixes for the low pollution (left panel), the moderate pollution (middle panel), and the high pollution scenarios (right panel), respectively.](image)

...and the error covariance matrices (horizontal bars) for the low emission (left panel), the moderate emission (middle panel), and the high emission levels (right panel), respectively. The high emission range was defined by profiles with Volume Mixing Ratios (VMRs) greater than or...
equal to 5 parts-per-billion-volume (ppbv) at surface. The moderate emission range includes the profiles with surface VMRs greater than or equal to 1 ppbv but less than 5 ppbv, or greater than 1 ppbv at any level between the surface and 500 hPa. The low emission is then defined as being lower than the lower bounds of the moderate emission range. The profiles were adjusted to match AIRS forward model levels. The modeled profiles are extrapolated near the surface with additional constraints to reflect values that are likely seen by satellite sensors.

Although for each pixel there are three possible a priori levels, the same set of the three-level a priori is used globally and throughout the AIRS data record. Thus, any spatial and temporal NH$_3$ variations detected using this algorithm are from AIRS measurements.

Figure 4. Correlation between the DBTI (difference of brightness temperature index) and DOFS (degrees of freedom for signal) for the three emission scenarios with low pollution in blue, moderate pollution in green, and high pollution in red.

To select one of the three a priori for each AIRS pixel, we examine the brightness temperature difference between a strong and a weak channel, divided by the measurement noise of the strong channel, defined as a “difference of brightness temperature index” (DBTI). This is similar to the method used by TES NH$_3$ and described by Shephard et al. (2011). The DBTIs vary with meteorological conditions and, most importantly, the thermal contrast at the surface. To take into
account of these effects, we simulate the relationship between the brightness temperature differences and TC under various meteorological conditions using SARTA. We randomly picked 13790 profiles from AIRS L2 V5 products over land from the months of January, April, July, and October in years 2003, 2008, and 2011. We then perturbed the NH\textsubscript{3} values spanning the three a priori mean profiles using the range of 0–100 ppbv for each atmospheric profile. The observed brightness temperatures are compared with the simulated values at a given thermal contrast (TC) to determine the level of a priori for the full retrievals. Figure 4 depicts a relationship between the DBTI and DOFS for the three emission levels with low emissions in blue, moderate emissions in green, and high emissions in red. The higher DBTIs are correlated with higher DOFS, which represent higher surface thermal contrast (Deeter et al., 2007).

4 QUALITY ASSURANCES AND ERROR ESTIMATES USED IN VERSION 3

4.1 Quality Assurance Levels

The NH\textsubscript{3} retrieval quality assurance levels are determined based on the retrieval sensitivities under various meteorological and surface conditions using the AKs and the DOFS. We also take into account the performance of the retrievals against surface thermal contrasts from AIRS products. Additionally, we examine the retrieval residuals, $\chi^2$, and the number of iterations to set proper quality assurance flags. The NH\textsubscript{3} retrieval quality is affected by the meteorological properties, such as the vertical temperature and water vapor profiles, surface temperatures, and emissivity, that are used to model the atmosphere. We also adapt the error information provided by the AIRS CCR for the relevant channels, which includes meteorological quantities that are used in deriving the AIRS CCR.
This error information is flagged by Q0, Q1, and Q2 with Q0 having the highest quality and Q2 being unusable. In the remaining discussions of this study, we used $\chi^2$ between 0.9 and 27, considering that the channels used are not all spectrally independent. The number of iterations limit was set at 10, meanwhile, only the cases with retrieval residuals less than 1 K are used. We also excluded cases with the surface thermal contrast between -4 and +4 K, to avoid ambiguous a priori levels; however, this primarily affects areas over the global oceans. Any additional screening of the data for higher quality requirements, e.g. the use of DOFS, will be discussed case by case. Although we have developed AIRS NH$_3$ products for all available datasets, only the daytime and land cases are discussed in this study. Additionally, only radiances with quality flag as Q0 are selected for the discussions in the following sections to ensure the best accuracy.

### 4.2 Error Estimates

The error estimates for this product are built on the available validation activities. Validations of retrievals using in situ measurements are vital to quantifying uncertainties in the concentrations, sources, transport patterns, and trends using satellite data. Direct measurements of tropospheric NH$_3$ are relatively sparse and in situ measurements above the ground level, necessary to validate satellite retrievals, are available for only limited locations and time periods (e.g. Nowak et al., 2007, 2010, 2012). Validation of AIRS NH$_3$ data sets with available in situ measurements is a continuous effort as more in situ measurements become available. Warner et al. (2016) used the DISCOVER-AQ NH$_3$ measurements over California (https://www-air.larc.nasa.gov/cgi-bin/ArcView/discover-aq.ca-2013). The sampling inlet and NH$_3$ calibration
set-up used during DISCOVER-AQ with the cavity ring down spectrometer (CRDS) (G2103, Picarro Inc.) is the same as used with the Chemical Ionization Mass Spectrometry (CIMS) and described in Nowak et al. (2007). The CRDS, aboard the NASA P-3B aircraft during DISCOVER-AQ Campaign, data period covers 16 January to 6 February 2013. The in situ NH$_3$ vertical profiles were made in the Southern San Joaquin Valley of California. This region inside the central valley of California, between the coastal mountains in the west and the Sierra Nevada Mountains in the east, consists largely of farmland with scattered dairy farms. Although most of the area is rural, the profiles were made near the small cities of Hanford and Corcoran. We only select spiral profiles from the flights within 45 km of the center of the retrieved AIRS profiles, for the closest match, and within 3 hours of the measurement window, similar to the method used for AIRS CO validation (Warner et al., 2006). Detailed validation procedures and figures can be found in Warner et al. (2016). We concluded that over regions with high NH$_3$ in situ concentrations, the convolved in situ profiles agree with the retrievals within > 1 to ~3 ppbv (~5–15 %) near the top of the boundary layer.

5 DATA SETS

5.1 Monthly Data Sets

These are gridded L3 monthly mean data sets. For file naming convention, file format and structure, data content, as well as options for reading the data and programming tools, see README Document for Global Ammonia Retrievals from AIRS (Advanced InfraRed Sounder) Satellite Measurements. For Dataset and Mission Instrument description, see User Guide for Global Ammonia Retrievals from AIRS (Advanced InfraRed Sounder) Satellite Measurements.
The data used in Warner et al. (2016; 2017) were from the V2 of this algorithm, and the current data have been updated using V3. The main difference is that the correlation between a layer above 500 hPa and a layer below is removed. This is because there is very little ammonia information above 500 hPa level in the current algorithm. We also updated AIRS L2 profiles noises but using the same V6 L2 profiles.

5.2 **Daily Data Sets**

The detailed description of AIRS NH$_3$ daily products will be added as the data sets become available, at a later time.

6 **REFERENCES**


Nowak, J.B., J. A. Neuman, K. Kozai, L. G. Huey, D. J. Tanner, J. S. Holloway, T. B. Ryerson, G. J. Frost, S. A. McKeen, and F. C. Fehsenfeld, “A chemical ionization mass spectrometry


AIRS Ammonia Algorithm Theoretical Basis Document Version 3.0


