Polar Radiant Energy in the Far-Infrared Experiment (PREFIRE) Algorithm Theoretical Basis Document (ATBD) for the 2B-ATM data product

Aronne Merrelli, Nathaniel B. Miller, Brian J. Drouin, Timothy I. Michaels, Erin Hokanson Wagner

R01 release, March/April 2025

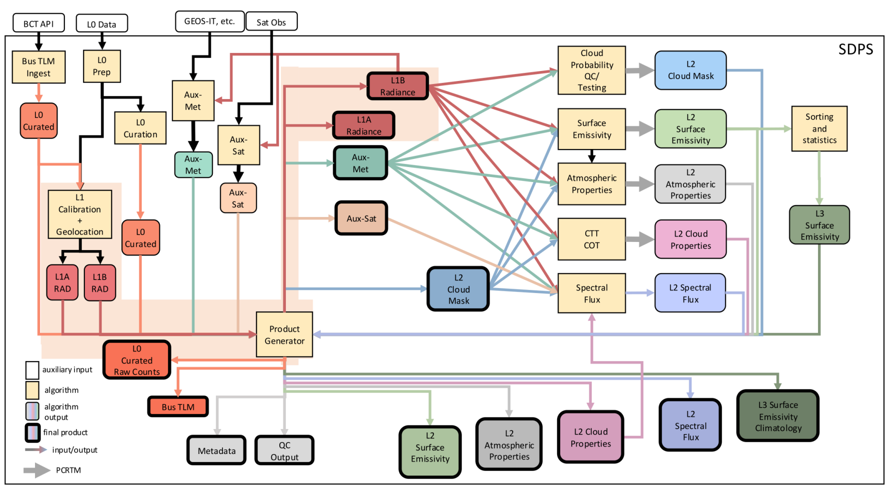


Figure 0.1: PREFIRE algorithm connectivity and flow.

Contents

1 Level-2 Atmospheric Retrieval

1.1 Introduction

This Algorithm Theoretical Basis Document (ATBD) describes the algorithms used to produce the 2B-ATM (clear-sky atmosphere) product for the Polar Radiant Energy in the Far Infrared Experiment (PREFIRE). The 2B-ATM algorithm uses data from the PREFIRE AUX-MET (Auxiliary Meteorological analysis), 2B-MSK (Cloud mask), and the 2B-SFC (surface properties) products as prior information.

1.2 Instrument overview

The spectrometer for PREFIRE, the Thermal Infrared Spectrometer (TIRS, or TIRS-PREFIRE), collects spectral radiance measurements across a wavelength range of approximately 5 to 54 μm with a spectral sampling of 0.84 μm. The light is dispersed by a grating onto a 64×8 element detector array that measures 8 simultaneous spectra along the spectrometer slit. The first four channels respond to shortwave radiance ($< 3\,\mu$m) and are not planned to be part of the calibrated 1B-RAD dataset, as there will be no calibration system for these wavelengths and no expectation of instrument performance. Due to the instrument design, there are two-channel gaps at approximately 7, 15, and 30 μm, at the boundaries of the order-sorting filters used to select for specific grating diffraction orders. The layout of the filters results in 54 usable channels covering most of the thermal infrared range. The actual flight detectors have individual bad detector elements which will imply a different number of valid channels between the 8 cross-track spectra. While the spectral resolution is much lower than modern infrared sounders, the spectral information available (particularly in the far-infrared water vapor rotational absorption band) does allow for coarse vertical resolution temperature and water vapor profiles.

Due to unpowered flight, the slowly decreasing altitude of the PREFIRE orbits will result in gradually smaller observational footprints throughout the mission. The initial ground footprint shapes of the 8 TIRS scenes are quadrilaterals, approximately 11.8 km x 34.8 km (cross-track by along-track) in size, with the 8 scenes separated cross-track by 24.2 km gaps between them. The temporal sampling rate of TIRS (0.7007 s) at the initial orbit altitude results in an along-track translation of only about 5.3 km, so that more than 6 consecutive measurements overlap. For the baseline 2B-CLD algorithm, no attempt is made to combine these observations in any way. In other words, each spectrum is treated as an entirely independent measurement. Future research will investigate whether the overlapping measurements can be combined in some way to reduce sensor noise.

Figure 1.1 shows a summary of the spectral response functions (SRFs) compared to a clear-sky, standard atmosphere emission spectrum. The SRFs are grouped according to the order sorting filter. Note the two channels (n) in the gaps between each grouping (n = 8, 9, 17, 18, 35, 36). Figure 1.2 shows sample weighting functions, for a standard atmosphere, across the primary wavelengths used for the 2B-ATM retrieval.

1.3 Overview

The 2B-ATM algorithm is a physical retrieval implemented with a standard optimal estimation approach (Rodgers, 2000), with a Levenberg-Marquardt parameter to adjust the weighting of the *a priori* and measurement information during iteration. The state vector consists of the temperature and water vapor vertical profiles and the surface temperature. The remaining relevant geophysical properties are taken from the values in the *a priori* datasets, and are assumed to be fixed values. The 2B-ATM algorithm is intended to be run in clear-sky conditions only (i.e., only on the measurements identified as clear by the 2B-MSK product).

A priori information is derived from several sources in the PREFIRE SDPS processing chain and input into the 2B-ATM algorithm, in order to create a final output data file. The output data file contains the final retrieval state at the coarse vertical resolution used in the 2B-ATM output product. Figure 1.3 shows the overall data flow.

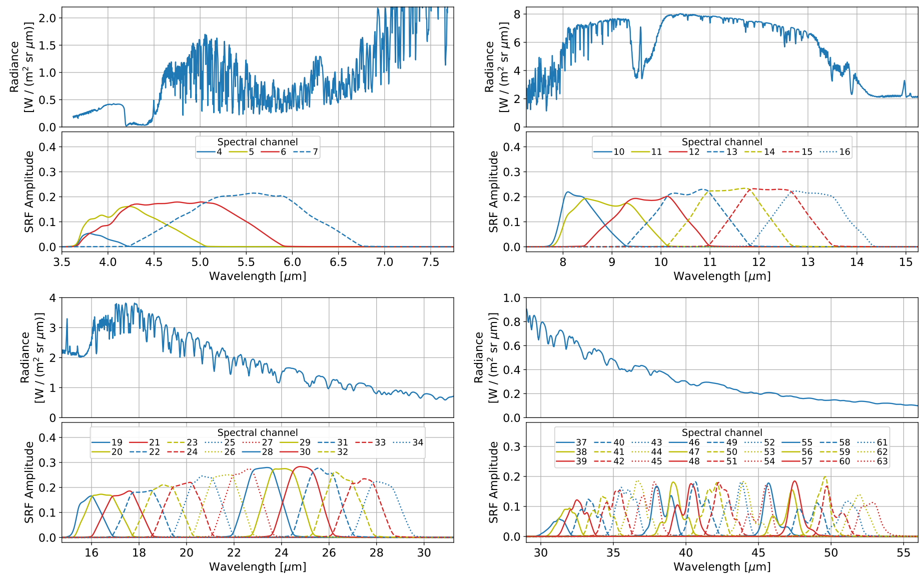


Figure 1.1: The TIRS SRFs grouped by the order sorting filter. In each panel, the upper plot is an emission spectrum from a standard atmosphere for comparison.

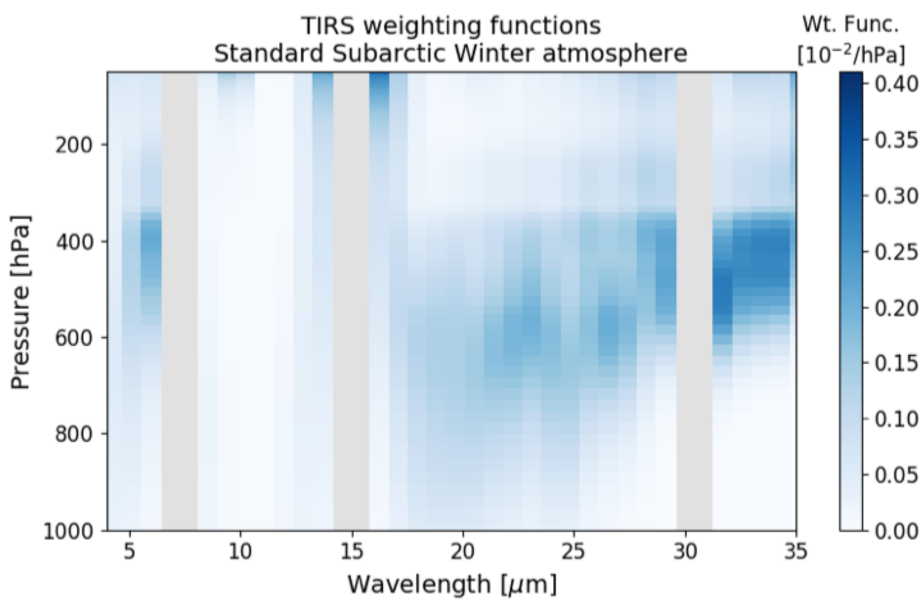


Figure 1.2: Sample TIRS weighting functions for the standard subarctic winter atmosphere, for $n = 4 - 40$.

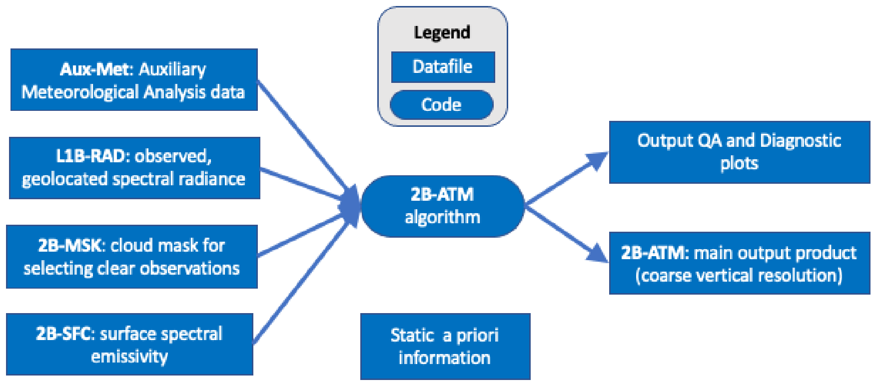


Figure 1.3: Algorithm flowchart

1.4 Spectral radiance forward model

The forward model used within the retrieval algorithm is the Principal Components-based Radiative Transfer Model (PCRTM) V3.4 (Liu et al., 2006). The PCRTM is an efficient and accurate plane parallel RT model, and uses a set of pre-computed Principal Components (PCs) describing a specific spectrometer sampling grid. The V3.4 implementation supports user defined profiles for the six primary infrared active molecules: water vapor, CO_2, O_3, CH_4, CO and N_2O. A number of other trace gases (such as CFCs) are included with fixed concentration profiles. The PCRTM can compute both the forward modeled radiance, as well as Jacobians for surface temperature, temperature profile, and concentration profiles for the six variable absorbers. A standard set of 101 fixed pressure levels defines the internal leveling grid for the forward model. To cover the FIR wavelengths measured by TIRS, we use a set of pre-computed coefficients constructed for a theoretical interferometer covering the wavenumber range $50 - 2760\,\mathrm{cm}^{-1}$ at a $0.5\,\mathrm{cm}^{-1}$ sampling grid. The high spectral resolution forward-modeled spectra are converted to a wavelength grid and the TIRS channel Spectral Response Functions (SRF) are applied to generate TIRS channel radiances.

For the 2B-ATM algorithm, the PCRTM is operated in a clear-sky mode, though the PCRTM does contain ice and water cloud spectral emissivity models. This version of the PCRTM does not include capability to model the spectral reflectance, so the output will include only the thermal emission.

1.5 State and measurement vectors

The retrieved state vector for the 2B-ATM algorithm includes the temperature profile, $\mathbf{T}$, the logarithm of the water vapor mass mixing ratio profile, $\ln(\mathbf{Q})$, and the surface temperature T_s, combined into a single joint as follows:

$$\mathbf{x} = [\mathbf{T}; \ln(\mathbf{Q}); T_s] \tag{1.1}$$

The profile variables are defined in the full PCRTM vertical level resolution, but for only the levels, j, with pressures less than the surface pressure as defined by the PREFIRE AUX-MET product (in other words, only the levels above the surface topography are retrieved). This implies that the typical number of retrieved levels will be slightly less than 101, or as few as 80 for very high-altitude surface topography. The below-surface levels are set to copies of the lowest-altitude air temperature and water vapor mixing ratio. The below-surface values are input to PCRTM for the forward radiance calculation, but they are not included in the state vector – meaning that the values do not change during optimization. (Note that the first below-surface level helps define the temperature and gas concentration of the partial layer containing the surface, so the below-surface values do impact the modeled radiance).

The measurement vector is the measured spectral radiance, as described in 1.2, and TIRS spectra will contain 54 valid spectral channels, less any identified bad detector elements. The shortest wavelength channels are not planned to be used in the 2B-ATM algorithm, in order to limit the impact of scattered solar radiation on the algorithm and limit any day/night biases that would arise. Our forward model (PCRTM v3.4) does not model the spectral reflectance, so these channels at the short wavelength end would require significant extra modeling efforts to be utilized in the algorithm. In addition, the information content for water vapor and temperature profiling of these channels is a small fraction of the total, so removing these channels from the retrieval does not cause a significant performance degradation. If the detector at $n = 6$ is not flagged as poor quality, then we will we pay close attention to its behavior in the flight data as this channel would be sensitive to both scattered solar radiation and non-Local Thermodynamic Equilibrium (non-LTE) emission in the 4.3 μm CO_2 absorption band (DeSouza-Machado et al. 2007). In addition, the longest wavelength channels (wavelengths larger than 40 μm) have relatively low signal to noise, and it may be necessary to remove several of these channels depending on their behavior in flight.

1.6 *A priori* data

The OE algorithm requires *a priori* covariance matrices and mean values describing the expected probability distribution of the state vector before the measurement is examined. Since our *a priori* is derived from auxiliary meteorological analysis data (the PREFIRE AUX-MET data product), the *a priori* mean will be direct copies of analysis fields interpolated to the TIRS observation locations and times. The covariance should then represent our expectation of the probability distribution errors in the meteorological analysis fields relative to the true values. This error distribution should include error in the analysis field itself, as well as error incurred from the interpolation of the analysis grid and time to the actual observation location and time. In practice, this covariance is very difficult to accurately estimate. As an empirical approximation, we can compute distributions of analysis errors from comparing an analysis time step to the average of the bracketing time steps. For example, we can compare the field at 12 UTC to the average of the fields at 09 UTC and 15 UTC (a $\pm$ 3-hour interpolation), or 06 and 18 UTC (a $\pm$ 6-hour interpolation). The averaged field is a proxy for the interpolation that will be done between a set of two analysis time steps and the arbitrary TIRS measurement time. This procedure should capture that part of the interpolation error, although it will not include any spatial interpolation. The covariance of these differences should capture realistic vertical error correlations, assuming the atmospheric transport within the analysis data is realistic. However, since this is not comparing the analysis data to an actual independent "truth" dataset, this method is likely to be an underestimate of the true error.

The PREFIRE SDPS will initially produce AUX-MET products from the GEOS-IT analysis data stream from NASA GMAO (Lucchesi, 2015) which will be available during the PREFIRE mission. To compute our initial covariance matrices, we used the above procedure on GEOS5 FP-IT data, which was available during algorithm development, before GEOS-IT was available. This data source has a time step of 3 hours for the three-dimensional variables (the temperature and water vapor fields), so in practice the interpolation time window would be a maximum of $\pm$ 1.5 hours to the arbitrary TIRS measurement times. We chose to use the interpolation from the $\pm$ 6 hour interpolation window as a way to compute a conservative estimate of the *a priori* covariance. In other words, this procedure will generate larger covariances in order to capture some of the known additional error sources that are not captured by the method. Finally, we approximate the computed covariances with an autoregressive correlation model (Lerner et al. 2003). This covariance model uses an exponential correlation model, so the covariance between two levels i and j, with variances σ_i^2 and σ_j^2, pressures p_i and p_j and a correlation scale p_L is given by (following Lerner et al. 2003, equation 7):

$$S_a(i,j) = \sigma_i \sigma_j \exp[-|p_i - p_j|/p_L] \tag{1.2}$$

Because the information content of TIRS observations is very low for the upper atmosphere ($p < 100\,\text{hPa}$), we use a correlation scale length in pressure coordinates that enforces a high degree of correlation among these upper atmosphere levels. Different values are used for the correlation scales and variance in the lower and upper atmosphere, and a logistic function smooths the transition between the two regimes. Table 1.1 gives the values of variances and correlation scales in the upper and lower atmosphere used in the analytic covariance model. Figure 1.4 shows the shape of the correlation structure and variance for the temperature *a priori* matrix. The correlation matrix for $ln(Q)$, is the same, and no correlation is assumed between temperature and water vapor.

Parameter	Upper atmosphere	Lower atmosphere
Surface T variance		$(2.0\,\mathrm{K})^2$
T variance	$(0.5\,\mathrm{K})^2$	$(2.0\,\mathrm{K})^2$
$\ln(Q)$ variance	$(0.3)^2$	$(0.6)^2$
T correlation scale	50 hPa	100 hPa
$\ln(Q)$ correlation scale	50 hPa	100 hPa

Table 1.1: Parameters defining the *a priori* covariance matrices for temperature and water vapor.

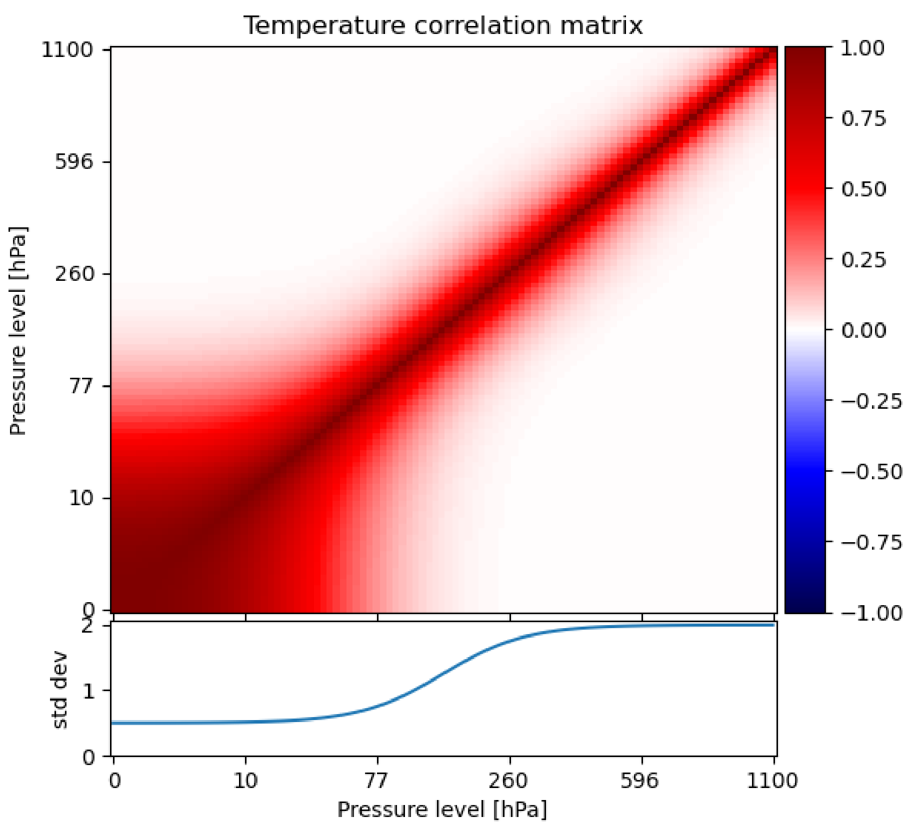


Figure 1.4: *A priori* temperature correlation and variance used in the retrieval.

In addition to the temperature and water vapor profiles that are present in the retrieval state vector, there are many additional geophysical variables that can impact the modeled radiance. These other variables have low information content in the TIRS measurements, and so cannot be retrieved, but should be specified accurately to minimize modeling errors. All of these additional variables are not included in the state vector, so their values will be fixed during optimization. These include the surface spectral emissivity and profiles of other infrared active molecules (CO_2, O_3, CH_4, CO, and N_2O). The surface spectral emissivity will be taken from the initial 2B-SFC retrieval. For CO and N_2O, the fixed profile from the standard atmosphere is assumed for all observations. Since the O_3 profile has strong vertical and temporal variation, it will be taken from the AUX-MET product, after passing through the same temporal and spatial interpolation as the temperature and water vapor profiles. For CO_2 and CH_4, a fixed volume mixing ratio profile is used, based on a climatology developed from the Copernicus Atmospheric Modeling System (CAMS) EGG4 greenhouse gas reanalysis product (Agustí-Pareneda et al., 2022). The climatology models the EGG4 XCO_2 and XCH_4 (the total column-averaged dry air mole fractions) as a function of latitude and time. To compute the climatology, the daily EGG4 data from 2003 - 2020 was first averaged across longitudes to produce zonal averages with time. The zonal averaged time series were then fit with harmonic series ($\sin + \cos$) and polynomial functions of time. These coefficients were then smoothed across latitudes with splines in order to reduce the overall dimensionality of climatology fits and to reduce noise across the individual latitude bands. Both CO_2 and CH_4 use linear polynomials. The CO_2 harmonic fit uses three terms, to fit annual and sub-annual cycles, while the CH_4 harmonic fit only uses one term. For each TIRS observation, the spline fits are evaluated at the observation latitude in order to determine the harmonic and polynomial coefficients, which are then evaluated to determine the CO_2 and CH_4 prior values. For an example with CO_2, assuming the spline fits yield the linear polynomial coefficients c_0, c_1 and the harmonic coefficients a_n, b_n, then the *a priori* CO_2 will be:

$$\mathrm{CO_2} = c_0 + c_1 t + \sum_{n=1}^{3} (a_n \sin(nt) + b_n \cos(nt)) \tag{1.3}$$

The evaluation for the CH_4 *a priori* value proceeds in a similar fashion, using fewer harmonic terms.

1.7 Inversion method

The inversion method used in the 2B-ATM algorithm is a standard Bayesian non-linear optimal estimation (OE) approach. Starting from an initial guess, the algorithm iterates the state vector, $\mathbf{x}$ value, recomputing the forward modeled spectral radiance and Jacobians at each step. The state vector updates at each iteration are the standard linear cost-function minimization steps. The method is similar to the standard Newton's method, with an additional Levenberg-

Marquardt parameter to adjust the weighting of the *a priori* and measurement information during iteration (Rodgers 2000). This implementation closely follows the OE solver method used in the NASA OCO-2 L2 algorithm (Crisp et al., 2021).

The cost function is from the standard OE formalism, following from an assumed *a priori* state vector mean ($\mathbf{x_a}$) and covariance (S_a), a measurement vector ($\mathbf{y}$) and measurement error covariance ($\mathbf{S}_\epsilon$), and a forward model function ($\mathcal{F}$). For a particular iteration where the state vector value is $\mathbf{x_i}$, the cost function (c) is given by:

$$c = (\mathbf{y} - \mathcal{F}(\mathbf{x_i}))^T \mathbf{S}_\epsilon{}^{-1} (\mathbf{y} - \mathcal{F}(\mathbf{x_i})) + (\mathbf{x_i} - \mathbf{x_a})^T \mathbf{S_a}{}^{-1} (\mathbf{x_i} - \mathbf{x_a}) \tag{1.4}$$

At each iteration, the forward model returns the modeled measurement ($\mathcal{F}(\mathbf{x_i})$) as well as the Jacobian at the state vector value ($\mathbf{K_i}$). These are used to compute the state update, $\mathbf{dx_{i+1}}$. The actual state update is computed using a linear matrix solver (the `linalg.solve` function in NumPy, which utilizes LAPACK). The state update equation is given by:

$$[(1+\gamma)\mathbf{S_a}{}^{-1} + \mathbf{K_i}{}^T \mathbf{S}_\epsilon{}^{-1} \mathbf{K_i}]\mathbf{dx_{i+1}} = [\mathbf{K_i}{}^T \mathbf{S}_\epsilon{}^{-1} (\mathbf{y} - \mathcal{F}(\mathbf{x_i})) + \mathbf{S_a}{}^{-1} (\mathbf{x_i} - \mathbf{x_a})] \tag{1.5}$$

These quantities are computed for a scaled state vector, using the matrix $\mathcal{M}$ which contains the inverse square root of the diagonal of $\mathbf{S_a}$ along the diagonal. Multiplying by $\mathcal{M}$ scales the vector by dividing by the per-state variable standard deviation contained in the $\mathbf{S_a}$ matrix. This yields a slightly different form of equation 1.5, operating in the scaled state vector space:

$$\mathcal{M} = \begin{bmatrix} S_{a,(1,1)} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & S_{a,(k,k)} \end{bmatrix}^{1/2} \tag{1.6}$$

$$\tilde{\mathbf{S}}_\mathbf{a} = \mathcal{M}^{-1} \mathbf{S_a} \mathcal{M}^{-1} \tag{1.7}$$

$$\tilde{\mathbf{x}} = \mathcal{M}^{-1} \mathbf{x} \tag{1.8}$$

$$[(1+\gamma)\tilde{\mathbf{S}}_\mathbf{a}^{-1} + \mathcal{M}\mathbf{K_i}{}^T \mathbf{S}_\epsilon{}^{-1} \mathbf{K_i}\mathcal{M}]d\tilde{\mathbf{x}}_\mathbf{i+1} = [\mathcal{M}\mathbf{K_i}{}^T \mathbf{S}_\epsilon{}^{-1} (\mathbf{y} - \mathcal{F}(\mathbf{x_i})) + \tilde{\mathbf{S}}_\mathbf{a}^{-1} (\tilde{\mathbf{x}}_\mathbf{a} - \tilde{\mathbf{x}}_\mathbf{i})] \tag{1.9}$$

After each iteration, the linearity of the forward model and cost function are assessed by comparing the value of the cost function after the state update (c_i+1) with a forecasted value of the cost function, assuming linearity, from the previous state value. The forecast value of the cost function relies on linear prediction of $\mathbf{y}$ at the current state:

$$\mathcal{F}(\mathbf{x_{i+1}}) \approx \mathcal{F}(\mathbf{x_i}) + \mathbf{K_i}\mathbf{dx_{i+1}} \tag{1.10}$$

The cost function forecast, c_{FC} is computed from equation 1.4, using equation 1.10 for $\mathcal{F}(\mathbf{x_{i+1}})$ and $\mathbf{x_{i+1}} \approx \mathbf{x_i} + \mathbf{dx_{i+1}}$. The ratio between the actual cost function change and the forecast cost function change, R, is calculated as:

$$R = \frac{(c_i - c_{i+1})}{(c_i - c_{FC,i+1})} \tag{1.11}$$

For a well-behaved update, the cost function at iteration i+1 should be smaller than i, and if it is linear then it should be close to the value of the forecast. In this case the ratio R should be close to 1. On the other extreme, for poorly-behaved updates, the cost function value may not change ($R = 0$) or may even increase after the update ($R < 0$). Given these limiting values for R, we assign one of four labels for the state update: divergent updates ($R <$ 0.0001), moderately nonlinear updates ($0.0001 < R < 0.25$), weakly nonlinear ($0.25 < R < 0.75$), or linear updates ($R > 0.75$). During iteration, a divergent update triggers a change to the Levenberg-Marquardt parameter (λ_0; see below), and the state update is discarded and recomputed with a new value of the λ_0 parameter. All other update types ($R > 0.0001$) will update the state vector using $\mathbf{dx_{i+1}}$.

The Levenberg-Marquardt λ_0 parameter controls a relative weighting between the *a priori* state estimate and the measurement. We start with a value of 10, and then update it during iteration with the following criteria, based on the classification of the state update by the cost function prediction method described above: For divergent updates or moderately nonlinear updates, λ_0 is increased by a factor of 10; for weakly nonlinear updates, λ_0 is unchanged, and finally, for linear updates, λ_0 is reduced by a factor of 2. The iteration is stopped when one of three exit criteria is reached: 1) the maximum number of iterations was reached, 2) the maximum number of divergent updates was reached, or 3) the state update was smaller than a threshold value. Retrievals that stop iteration by the third criterion are considered converged. The third criterion is evaluated by computing the size of the squared scaled state vector update relative to the current posterior error covariance, divided by the number of state vector elements. This quantity (z) is compared against a threshold value of 0.1 to determine convergence:

$$z = \frac{1}{k}\mathbf{d\tilde{x}^T_{i+1}}\mathbf{\tilde{S}}^{-1}\mathbf{d\tilde{x}_{i+1}} \tag{1.12}$$

Each one of the criteria is set by an adjustable parameter that will be continually re-evaluated. Each of the thresholds can represent tradeoffs between algorithm throughput, yield, and accuracy, that will not be fully characterizable with pre-flight simulation testing. For example, more retrievals will converge if the iteration limit is increased, or higher accuracy might be obtained with smaller state update thresholds.

Finally, when the iterative algorithm stops, we perform a final forward model calculation to get the value of $\mathcal{F}(\mathbf{x}')$ at the retrieved state for the purposes of a χ^2 calculation to evaluate the goodness of fit and to compute the spectral residuals. The χ^2 is computed in the conventional way, given by equation 1.13.

The number of degrees of freedom are taken from the trace of the averaging kernel matrix, $\mathbf{A}$, at the final state vector value. This allows for the calculation of the reduced χ^2, which is expected to be near a value of 1 for a successful retrieval. In the following equation, k is the number of variables in the state vector, and d is the number of degrees of freedom in the retrieval.

$$\chi_\nu^2 = \frac{(\mathbf{y} - \mathcal{F}(\tilde{\mathbf{x}}))\mathbf{S}_\epsilon{}^{-1}(\mathbf{y} - \mathcal{F}(\tilde{\mathbf{x}}))}{k - d} \tag{1.13}$$

While the actual measurement vector used in the retrieval will be a subset of the total TIRS channels, the final forward model calculation will produce modeled radiances for all 54 valid channels. The full spectral model will be used for quality assessment as described in the section below.

1.8 Output processing

1.8.1 Layer specification of output profiles

Because the information content of the TIRS measurement is relatively low, the full vertical resolution profiles (the PCRTM standard 101 levels) are far more densely spaced than what is needed to capture the actual information content. This is particularly true in the upper atmosphere levels, where the full resolution includes 44 levels below 100 hPa. Before creating the final output product, the full resolution levels are grouped and combined into a smaller number of layers. A general heuristic is used to determine an appropriate grouping of the high vertical resolution levels based on the information content profile at the high vertical resolution. Starting from the TOA level, the information content profile is integrated downwards (with a cumulative sum), stopping when the total information reaches some threshold level. The threshold level should be less than one, otherwise the information content total would suggest the retrieval would have enough degrees of freedom for signal to retrieve partially independent sub-layers within the layer. Once the threshold level is surpassed, the level defines the bottom level that will be used in the combined layer. The process is repeated with the current level as the top level of the next combined layer.

The level combination process produces different level groupings for each individual profile. In particular, the process would produce a smaller number of combined layers for low information content profiles (high surface altitude, dry conditions). We desired a single layer specification that is applicable globally. Therefore, we determined the layer specification from an ensemble of tropical ocean profiles from ERA5 reanalysis data, as these will tend to have the highest total information content, and we determined a single set that is approximately the number returned from the average temperature profile. For simplicity, the same layer specification is applied to both the temperature and water vapor profile.

When the high-resolution profiles are combined, each low vertical resolution layer value is the mean of the associated high-resolution grouping. The posterior covariance matrix is computed by block averaging according to the same layer

Combined layer number	PCRTM level number range	Pressure range (hPa)	Pressure thickness (hPa)
1	1 – 51	0.005 – 156	156
2	52 – 64	156 – 307	151
3	65 – 72	307 – 433	126
4	73 – 79	433 – 565	132
5	80 – 86	565 – 718	153
6	87 – 93	718 – 892	174
7	94 – 101	892 – 1100	208

Table 1.2: Specification of the combined layers for the output product.

specification. The layer specification is shown in Table 1.2. The pressure boundaries of the coarse output levels are specified as the half-levels between the lower and upper pressure levels of the neighboring combined layers. Note that nearly the entire upper atmosphere ($p < 150\,\mathrm{hPa}$) is combined into one coarse output layer. The troposphere is then divided into six layers with pressure thicknesses ranging from approximately 120 – 170 hPa.

1.8.2 Quality assessment variables

The optimal estimation algorithm outputs several status variables that are used for quality assessment. These are combined into two quality variables. First, a summary integer flag (`atm_quality_flag`) with categorical values is provided. A second bit flag variable (`atm_qc_bitflags`) that includes more detailed status information is included in the product. The summary integer flag should be sufficient for most uses, with the additional detail in the bit flags available for more advanced analysis of the product. The details of the flags will likely change with on-orbit data, but the intention is for the summary integer flag to retain the same definition, while the bit flags will likely include additional status conditions as needed.

The summary integer flag records the four main status conditions for the 2B-ATM algorithm within each TIRS observation. In brief, these are that no retrieval was attempted, or that a retrieval was attempted resulting in one of three outcomes: a good-quality converged retrieval, a poorer-quality converged retrieval, or an unconverged retrieval. These are listed in table 1.3 with the corresponding integer values. The bit flags give more detail about the iteration convergence, which can fail because either the diverging step or iteration count limits were reached, or an unphysical state vector value was reached during iteration. Additional status bits may be added once on-orbit data is analyzed. Table 1.4 describes the current set of bit flags within the algorithm.

Threshold values for these QA variables are still under assessment using simulated data, and they will be re-evaluated as further refinements of the preflight noise models from TIRS become available. Ultimately, the thresholds will be reassessed again when flight data becomes available.

status value	Description
0	Best quality, converged retrieval
1	Poor quality, converged retrieval (reduced χ^2 exceeds threshold)
2	Retrieval did not converge

Table 1.3: Description of 2B-ATM integer quality flag.

bit number	Status description
0	reduced χ^2 threshold exceeded
1	retrieval exceeded iteration count limit
2	retrieval exceeded diverging step count limit
3	retrieval went outside allowable state vector range
4	retrieval solver crashed
5	retrieval used constant surface emissivity
10	retrieval not attempted due to cloud mask
11	retrieval not attempted due to latitude constraint
12	retrieval not attempted due to 1B-RAD status

Table 1.4: Descriptions of 2B-ATM quality bit flags.

As described in the earlier section, when the retrieval is performed, the forward model is used to create a full spectral model even if a subset of the channels was used in the retrieval. If there are detector elements that have unexpected behavior on orbit, the modeled radiance may allow for assessment of the radiometric response relative to the "good" detector elements. For example, a problematic channel could be evaluated by excluding it from the retrieval and analyzing the residuals (modeled minus observed radiance) in the retrieval output. These spectral residuals will be continually monitored during the mission. Furthermore, bulk statistics based on the fraction of converged retrievals and the averaged reduced χ^2 will be used to assess overall retrieval product quality. These values will be visualized separately for the eight TIRS spectra per instrument. Differences between the average performance among the eight spectra (for example, the distribution of χ^2 values) can indicate problems with the 1B-RAD calibration.

1.9 Retrieval analysis: information content and uncertainty characterization

Tests were performed with simulated measurements in order to assess the performance and uncertainties computed by the retrieval algorithm. The simulated measurements were generated from radiance simulations using the same forward model used in the retrieval (PCRTM, as described earlier), based on atmospheric profiles from ERA5 reanalysis data (Hersbach et al. 2020) produced by the European Centre for Medium-range Weather Forecasts. The region locations reflect

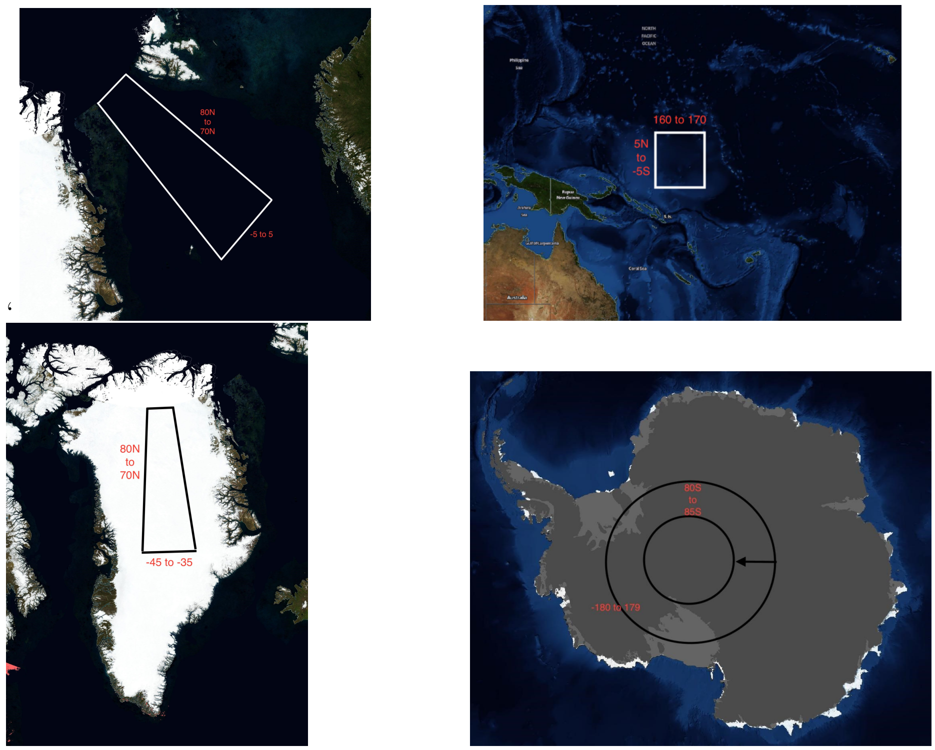
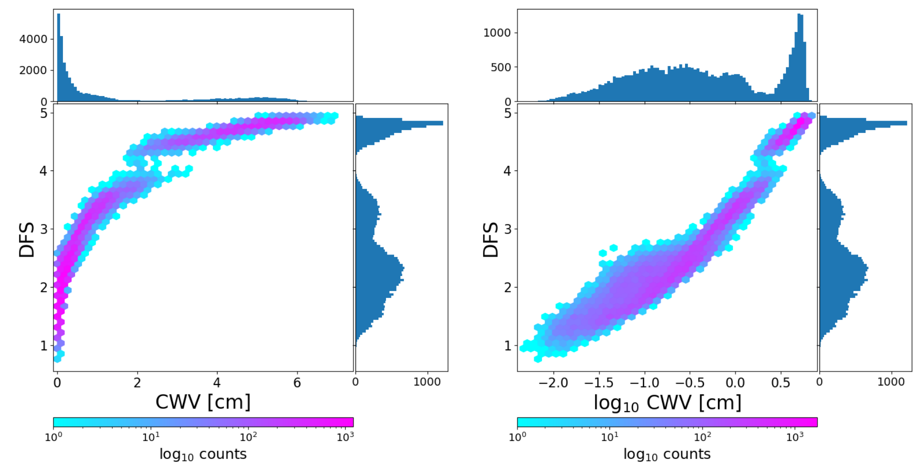


Figure 1.5: Regions where profile ensembles were extracted from ERA5 reanalysis data: Upper left, Arctic ocean (70^oN – 80^oN, 5^oW – 5^oE); upper right, tropical ocean (5^oS – 5^oN, 160^oW – 170^oE); lower left, Greenland ice sheet (70^oN – 80^oN, 45^oW – 35^oE); Antarctica (80^oS – 85^oS). Images are from NASA Worldview.

that PREFIRE is a polar-focused mission but will also produce global products. Three of the regions are polar (Arctic ocean, Greenland ice sheet, Antarctica) and the fourth region is near the warm pool in the tropical Pacific Ocean. The full set of profiles thus spans a wide range of polar climate conditions, and also spans the range of cold/dry and warm/moist extremes we expect to observe globally. Figure 1.5 shows the locations of the four regions. For each region, we drew 8000 random samples from 2016 to yield a testing ensemble. Each profile produces a simulated radiance and Jacobian from PCRTM, which is then processed through the standard "Degrees of Freedom for Signal" (DFS) analysis (Rodgers, 2000) to assess the information content of the temperature and water vapor profiles. The TIRS information content is very strongly modified by the total water vapor amount, since the temperature information is relatively low because of the masked channels at 15 μm. Figure 1.6 shows a joint histogram of the DFS and total column water vapor (CWV) across the entire 32000 profile ensemble.



(a) Linear x-scale for CWV. (b) Logarithmic x-scale for CWV.

Figure 1.6: Joint and marginal histograms of retrieval DFS (for temperature and water vapor profiles) versus a linear or logarithmic scaled total column water vapor (CWV). The data from the three polar regions populates a similar region in these two parameters, with the tropical data separated to larger values in both DFS and PWV.

The profiles are used to directly generate simulated observations with representative sensor noise, but then the retrieval is run with a perturbed profile as the first guess in conjunction with these simulated observations. The perturbation is created as a correlated random variable, drawn from a covariance nearly equal to the prior. The same correlation structure is used, but the variance is constant with height rather than decreasing to a lower value in the upper atmosphere.

The retrieval is run on all selected profiles, and the statistics are pooled within each regional ensemble. The difference between retrieved and true profiles is characterized in terms of the precision (standard deviation of the differences) and accuracy or bias (mean of the differences). Furthermore, we examine the accuracy of the retrieval algorithm's reported uncertainty by computing the standard deviation of the scaled differences (z):

$$z = (\hat{\mathbf{x}} - \mathbf{x}_{\mathbf{true}})/\epsilon_x \tag{1.14}$$

where the uncertainty (ϵ_x) is the posterior value computed by the retrieval algorithm. With accurate uncertainties, this quantity should be a normally distributed value with unit variance, meaning the standard deviation should converge to 1.

Figure 1.7 shows the results for the temperature profiles from the Arctic ocean ensemble. The results from the other three regional ensembles are similar and omitted for brevity. The results are generally within expectations. In gen-

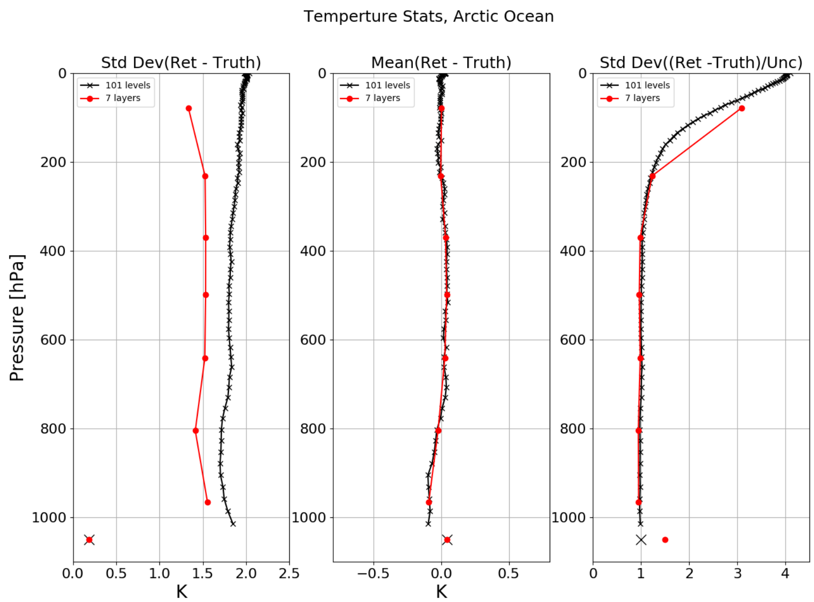


Figure 1.7: Evaluation statistics for temperature for the Arctic ocean ensemble: (a) retrieval precision estimate, from the standard deviation of retrieval – truth differences; (b) retrieval accuracy estimate, from the mean of the differences; (c) accuracy of uncertainties, from the standard deviation of the scaled uncertainties. The original 101-level and combined 7-layer profiles are shown, with the surface temperature displayed as the separated single point at p=1050 hPa.

eral, within the troposphere the reduction in uncertainty relative to the prior in the full set of retrieval levels ranges from 0.1 to 0.4 K. The level combining does increase the precision slightly. For example, the standard deviation in Figure 1.7(a) drops from about 1.8 to 1.5 K, due to averaging out of uncorrelated error among the combined levels. The biases (middle column) are small, generally 0.1 K or less, with the larger biases generally occurring near the surface. Note that the scaled uncertainty is very close to 1 for the original and combined levels throughout the troposphere, demonstrating that the post-processing calculations are handling the uncertainty propagation accurately through the level combining process. The scaled uncertainty increases to larger values in the upper atmosphere, here due to the perturbations being larger than the *a priori* assumptions. The measurements also have very little information about the upper atmosphere levels (note in 1.7(a) that the standard deviation does not drop from the perturbation amplitude of 2 K), so the retrieved state is very close to the prior.

Figure 1.8 shows the results for the water vapor profiles, displayed in a similar way as the temperature results in Figure 1.7. Since the water vapor

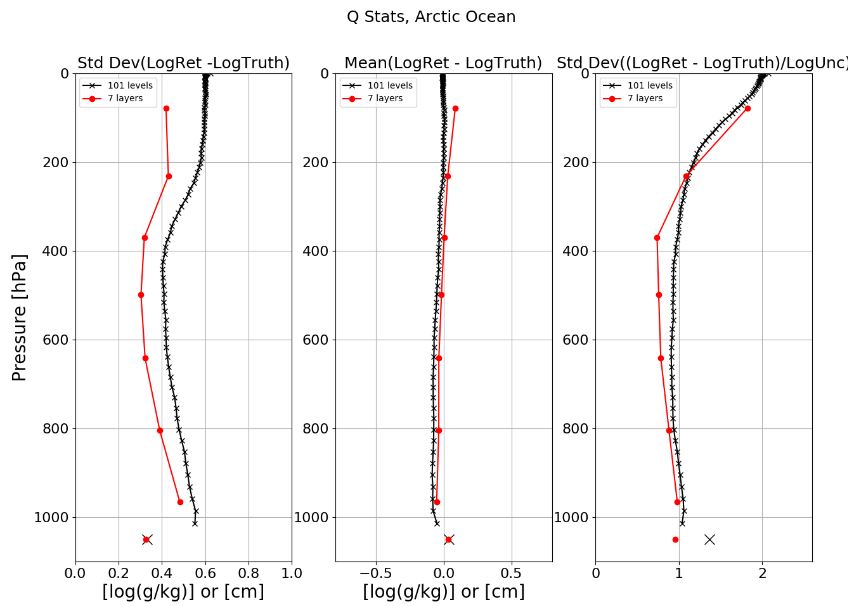


Figure 1.8: Water vapor evaluation statistics. Similar to Figure 1.7 (for temperature) but using the retrieval statistics in $\ln(Q)$ space. The separated point at p=1050 hPa is the results for the CWV.

profile retrieval is done in $\ln(Q)$ space, the accuracy and precision estimates are also shown in the $\ln(Q)$ space. The transformation back to linear space will introduce a bias itself, so it is important to isolate any possible bias from the retrieval itself. Displaying the results in $\ln(Q)$ space also makes it easy to relate to the variance of the *a priori*. The results for water vapor are largely the same as temperature, and there is again good consistency between the regional ensembles.

The primary output variable for the retrieval is the column water vapor (CWV). The CWV is simply the vertical integral of the water vapor profile, but it is important to ensure the uncertainty propagation through the integration is accurate. In Figure 1.8(c), the scaled uncertainties for the CWV computed from the original 101-level profile are slightly high, but after level combination the CWV integral uncertainty is very accurate. Table 1.5 lists the final CWV uncertainties observed across the four regional ensembles and gives a rough estimate of the overall uncertainty of global CWV estimates from the 2B-ATM retrieval.

	Antarctica	Greenland	Arctic ocean	Tropical ocean
Mean CWV [cm]	0.111	0.201	0.880	4.92
Standard dev. [cm]	0.056	0.082	0.327	0.89
Fractional uncertainty (Std. dev./mean)	51%	41%	37%	18%

Table 1.5: Uncertainty estimates of column water vapor from the ATM retrieval.

1.10 Validation plan

Comparisons of ground-based and airborne measurements to PREFIRE water vapor level 2 products provide an independent gauge of the accuracy of the retrievals. Validation sources are critical to verifying the retrieved values are within the reported uncertainty. Ground-based stations, situated in the Arctic and Antarctica, have a lot of variation of data availability and measurement types. Direct measurements of the temperature and humidity profiles throughout the troposphere via radiosondes are extremely valuable, but usually occur only once or twice a day at a given station. Sub-setting PREFIRE overpasses with station locations and known launch times throughout the mission provides a repository of data useful for statistical analysis. We will utilize the Integrated Global Radiosonde Archive (IGRA) v2 (Durre et al. 2006) which includes a consolidated distribution of radiosonde observations (ftp://ftp.ncdc.noaa.gov/pub/data/igra). Stations in Antarctica and the Arctic are most pertinent for the validation of PREFIRE products. In Antarctica, a majority of active stations in the IGRA are located along the coast as seen in Figure 1.9. The 14 stations labeled in Figure 1.9 will be used during clear-sky scenes to compare retrieved specific humidity profiles to direct measurements from radiosondes. In addition, column water vapor (CWV) estimates will be compared to the integrated water vapor from the radiosondes. The station located at the South Pole is not included as that latitude falls outside of the orbital range of PREFIRE. Specific station data usage is subject to data quality-control tests and availability over the mission period.

In the Arctic, there are 63 currently active stations in the IGRA that average one or more radiosonde launches per day and are north of 60^{o}N (Figure 1.10). Similar analysis to the Antarctic sites will be performed by comparing the Northern hemisphere sites to the subsampled PREFIRE data. In addition, the three sites labeled in Figure 1.10 have instruments designed to measure additional atmospheric state information, including high temporal resolution PWV estimates from microwave radiometers and various cloud properties. Specific station data usage is subject to data quality control tests and availability over the mission period.

The Simultaneous Nadir Overpass (SNO) method allows for comparison between PREFIRE and other polar-orbiting satellites (e.g., Cao et al. 2005) when the two similar nadir overpasses occur within small time difference. The Cross-track Infrared Sounder (CrIS) and Infrared Atmospheric Sounding Interferome-

Figure 1.9: IGRA station locations around Antarctica.

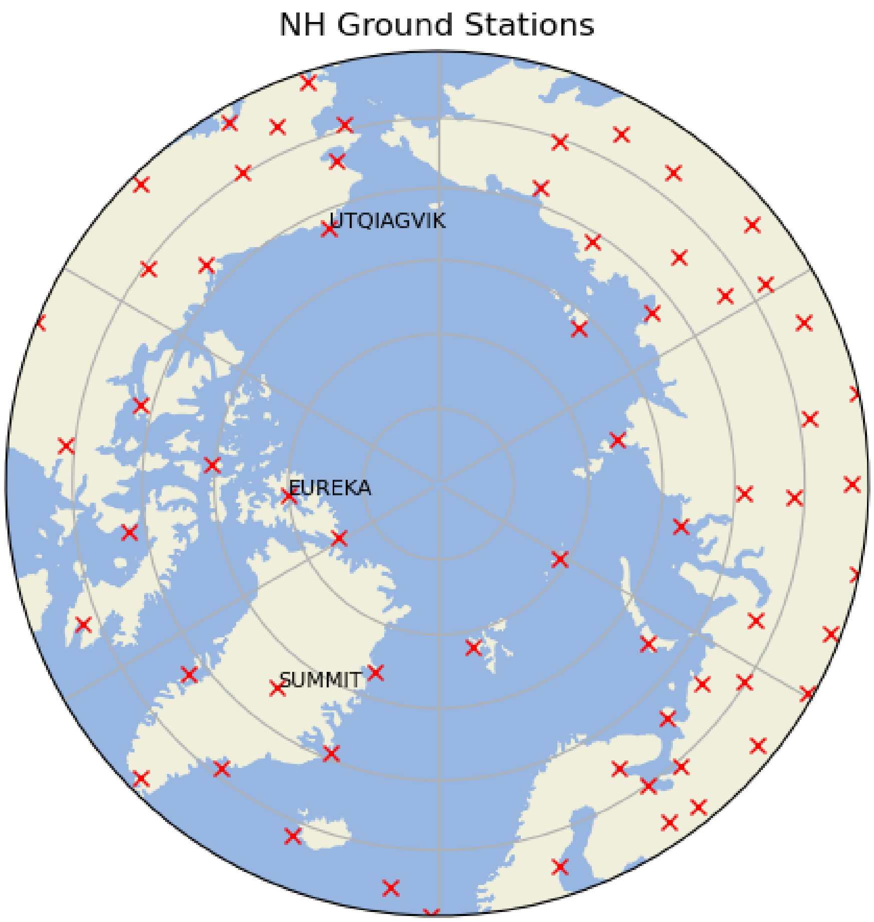


Figure 1.10: IGRA station locations around the Arctic Ocean.

ter (IASI) derived products provide the ability to directly compare water vapor retrievals via SNOs. These comparisons of similar Earth scenes may bring to light potential calibration issues, possible biases in water vapor retrievals, or scene-dependent discrepancies.

1.11 References

Agustí-Panareda, A., Barré, J., Massart, S., Inness, A., Aben, I., et al.: Technical note: The CAMS greenhouse gas reanalysis from 2003 to 2020, Atmos. Chem. Phys., 23, 3829–3859, https://doi.org/10.5194/acp-23-3829-2023, 2023.

Cao, C., C., H. Xu, J. Sullivan, L. Mcmillin, P. Ciren, and Y. Hou, 2005: Intersatellite radiance biases for the High Resolution Infrared Radiation Sounders (HIRS) on-board NOAA-15, -16, and -17 from simultaneous nadir observations. J. Atmos. and Ocn. Tech., 22, 381-395.

Crisp, D., O'Dell, C., Eldering, A., Fisher, B., Oyafuso, F., et al., 2021. OCO (Orbiting Carbon Observatory)-2 Level 2 Full Physics Retrieval Algorithm Theoretical Basis, Tech. Rep. OCO D-55207, NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, Version 3.0 - Rev 1, available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO_L2_ATBD.pdf (last access: Jun 4 2021),

Copernicus Atmosphere Monitoring Service: CAMS global greenhouse gas reanalysis (EGG4), CAMS Atmosphere Data Store (ADS) [data set], https://doi.org/10.24380/8fck-9w87, 2021.

DeSouza-Machado, S.G., Strow, L.L., Hannon, S.E., Motteler, H.E., Lopez-Puertas, M., Funke, B., Edwards, D.P., 2007. Fast forward radiative transfer modeling of 4.3 um nonlocal thermodynamic equilibrium effects for infrared temperature sounders. Geophys. Res. Lett. 34.

Durre, I., R. S. Vose, and D. B. Wuertz, 2006: Overview of the Integrated Global Radiosonde Archive. Journal of Climate, 19, 53-68.

GMAO website, last accessed 2022-06-24.https://gmao.gsfc.nasa.gov/GMAO_products/GEOS-IT/

Hersbach, H, Bell, B, Berrisford, P, et al. The ERA5 global reanalysis. Q J R Meteorol Soc. 2020; 146: 1999– 2049. https://doi.org/10.1002/qj.3803

Liu, X., Smith, W.L., Zhou, D.K., Larar, A., 2006. Principal component-based radiative transfer model for hyperspectral sensors: theoretical concept. Appl. Opt. 45, 201–209.

Lerner, J.A., Weisz, E., Kirchengast, G., 2002. Temperature and humidity retrieval from simulated Infrared Atmospheric Sounding Interferometer (IASI) measurements. J. Geophys. Res. 107, 4189. https://doi.org/10.1029/2001JD900254

Lucchesi, R., 2015: File Specification for GEOS-5 FP-IT. GMAO Office Note No. 2 (Version 1.4), 60 pp, available from http://gmao.gsfc.nasa.gov/pubs/office_notes .

Rodgers, C. (2000) Inverse Methods for Atmospheric Sounding: Theory and Practice. World Scientific Publishing Co Pte. Ltd.

2 Appendix

2.1 Table of variables and symbols

$\mathbf{A}$	averaging kernel matrix
α	angular resolution
β	azimuth angle
B	blackbody radiance
BW	spectral bandwidth
χ	convergence criterion
c	speed of light, cost function
CED	Cloud particle Effective Diameter
COD	Cloud Optical Depth
CTP	Cloud Top Pressure
CWP	Cloud Water Path
d	degree of freedom
ε	emissivity
ϵ	noise, error
ϕ	longitude
E	irradiance
F	flux
f	focal length
$\mathcal{F}$	function
γ	*a priori* weight
G	gravitational constant
g	gain
H	height
h	Planck's constant
I	radiance
IC	Information Content
IWC	Ice Water Content
IWP	Ice Water Path
j	counter
k	Boltzmann's constant, unknown
$\mathbf{K}$	Jacobian
λ	wavelength, Marquardt-Levenberg parameter
l	distance
L	radiance
LTS	Lower Tropospheric Stability
LWC	Liquid Water Content
LWP	Liquid Water Path
M	counter, mass
m	number of along-track frames
$\mathcal{M}$	matrix

N	counter
n	channel
$\mathcal{N}$	normal distribution
ν	frequency
NEdT	Noise-Equivalent delta Temperature
o	offset
Ω	solid angle
p	pressure
P	probability
PWV	Precipitable Water Vapor
Q	water vapor
ρ	reflection coefficient
R	radius, resistance, cost-function change
$\Re$	response function
$\wp$	responsivity
σ_B	Stefan-Boltzmann constant
$\mathbb{S}$	signal level in digitized counts
S	covariance
SI	Segmentation Index
SNR	Signal-to-Noise Ratio
SRF	Spectral Response Function
θ	latitude, potential temperature, polar coordinate angle
τ	transmission, optical depth
T	temperature
TR	Training Radiances
TREM	TRaining Eigenvector Matrices
t	time
ϕ	polar coordinate angle
V	voltage
v	velocity
x, y, z	position coordinates
z	convergence, standard deviation of scaled differences
$\mathbf{x}$	state vector
X	focal plane position
$\mathbf{y}$	measurement vector
Y	focal plane position
ζ	incidence angle

Table 2.1: Table of variables and symbols.

2.2 Abbreviations and acronyms

ADM	Angular Distribution Model
AIRS	Atmospheric Infrared Sounder
ATBD	Algorithm Theoretical Basis Document
CERES	Clouds and the Earth's Radiant Energy System
DEM	Digital Elevation Model
DOF	Degree of Freedom
ECI	Earth-Centered Inertial
ECMWF	European Centre for Medium-Range Weather Forecasts
EOF	Empirical Orthogonal Function
FIR	Far-InfraRed
FOV	Field Of View
FPA	Focal Plane Array
FWHM	Full Width at Half Maximum
GEOS-IT	Goddard Earth Observing System for Instrument Teams
GMAO	Global Modelling and Assimilation Office
IFOV	Instantaneous Field Of View
IFS	Integrated Forecasting System
LW	Longwave
MIR	Mid-InfraRed
NASA	National Aeronautics and Space Administration
NEP	Noise Equivalent Power
NEdR	Noise Equivalent delta spectral Radiance
OE	Optimal Estimation
OLR	Outgoing Longwave Radiation
PCRTM	Principal Component-based Radiative Transfer Model
PREFIRE	Polar Radiant Energy in the Far-InfraRed Experiment
ROIC	Read-Out Integrated Circuit
RMSE	Root Mean Square Error
SDPS	Science Data Processing System
SSF	Single Scanner Footprint
SRF	Spectral Response Function
TCWV	Total Column Water Vapor
TIRS	(TIRS-PREFIRE) Thermal InfraRed Spectrometer
TIRS1	Thermal InfraRed Spectrometer on PREFIRE-SAT1
TIRS2	Thermal InfraRed Spectrometer on PREFIRE-SAT2
TOA	Top of Atmosphere
UTC	Coordinated Universal Time
VZA	Viewing Zenith Angle
WV	Wavelength

Table 2.2: Abbreviations and acronyms.

2.3 Figure listing with links

Table 2.3: List of Figures in this ATBD.