# DEVELOPMENT OF A CLEAR-SKY ATMOSPHERIC RETRIEVAL ALGORITHM USING FAR INFRARED SPECTRAL MEASUREMENTS FOR PREFIRE

by

Meredith Grames

A thesis submitted in partial fulfillment of

The requirement for the degree of

Master of Science

(Atmospheric Science)

at the

UNIVERSITY OF WISCONSIN – MADISON

2022

# Abstract

The far infrared (FIR) radiation is important in Earth's climate system. It comprises nearly half of the radiation emitted by Earth's surface. It is also sensitive to changes in atmospheric water vapor. Measurements in the FIR are therefore, desirable in the study of Earth's changing climate. However, space-borne spectral measurements of the FIR have not been taken since the late 1970s. In recent years, thanks to renewed interest in the FIR and advances in remote sensing technology, several missions have been announced that will launch satellites that will take spectral measurements of the FIR once more. This paper discusses the development of a clear-sky atmospheric retrieval algorithm for the Polar Radiation in the Far InfraRed Experiment (PREFIRE). The algorithm regression coefficients taken from principal component regression to retrieve temperature and water vapor profiles. The algorithm is tested using simulated measurements from the Thermal Infrared Spectrometer (TIRS) developed for PREFIRE. In addition, we test the viability of Brightness Temperature Classification as a means of retrieval improvement. We also test the retrieval with and without FIR measurements to analyze the additional information that may be gained from FIR spectral measurements. Overall, we find that the TIRS retrieval is comparable to that of IASI, an instrument with a much higher spectral resolution that only measures in the midinfrared. Brightness Temperature Classification resulted in a reduction of retrieval error for temperature retrievals, and improved water vapor retrievals at lower altitudes. These changes were particularly strong in the Arctic. We also find that the inclusion of FIR spectral measurements in atmospheric retrievals can decrease temperature retrieval error by 20%, and water vapor retrieval error by 6%.

# Table of Contents

# Chapter 1: Background

## 1.1 The Far Infrared

The infrared portion of the spectrum is subdivided differently throughout literature and depending on the purpose of research. Some simply divide the infrared into shortwave infrared (having wavelengths shorter than about 4μm) and longwave or thermal infrared (wavelengths longer than 4μm). The thermal infrared may be further divided into the mid and far infrared. The definitions of the mid and far infrared vary. For this paper, we define the far infrared (FIR) as having a wavelength between 15 and 50 μm (or wavenumber from 667 to 200 $cm^{-1}$). Infrared wavelengths between about 4 and 15 μm are classified as mid infrared (MIR). This follows the definition found in Harries et al. (2008) and aligns well with current observing system capabilities since today's satellites only measure up to 15 μm.

### 1.1.1 Spectral Characteristics of the FIR

The far infrared is dominated by water vapor absorption. The water vapor rotational absorption band is located within the FIR part of the spectrum, and water vapor absorption is strong here due to the large number of strong absorption lines. These closely spaced strong absorption lines also produce a broad continuum absorption, which influences the atmospheric transmissions well away from the absorption line centers. This phenomenon was first noticed in the late 18 and early 1900s, where absorption of atmospheric radiation seemed to be occurring within atmospheric windows (Rubens & Aschkinass, 1898; Hettner, 1918, as cited in Shine, Ptashnik & Rädel, 2012). In the 1930s, the cause of this anomalous absorption was determined to be the far wings of other absorption bands (Elsasser, 1938, as cited in Shine, Ptashnik & Rädel, 2012). These bands are made of a large number of strong absorption lines, strong enough that the

wings cause noticeable absorption across the infrared, resulting in continuous absorption by water vapor, hence water vapor continuum (Shine, Ptashnik & Rädel, 2012).

In addition to water vapor, there is a $CO_2$ absorption band centered at 15μm (667 $cm^{-1}$) that extends to approximately 17 μm. Other atmospheric trace gases, such as $O_3$ and $N_2O$ do absorb in the far infrared, but their impact is small compared to absorption from water vapor and carbon dioxide (Harries et al., 2008).

### 1.1.2 The Importance of Water Vapor Measurements

The strong sensitivity of FIR radiation to water vapor makes these wavelengths particularly attractive for retrieving water vapor profiles in the atmosphere, as it is more sensitive to water vapor than the MIR. Water vapor plays a key role in both weather and climate. For example, our knowledge of the mechanics of climate change are related to a water vapor feedback effect: increases in other greenhouse gases, such as carbon dioxide, result in some heating that leads to more water evaporation from the surface, and thus more water vapor ending up in the atmosphere. Water vapor is the strongest of the greenhouse gases, but is more variable and has a shorter atmospheric lifetime than other gases, including $CO_2$. However, when more water vapor is in the atmosphere, warming increases, resulting in further evaporation from the surface, creating a positive feedback in the climate system.

The above explanation of the role of water vapor is very broad and simplified. There are many additional processes that relate to water vapor and the changing climate. However, these processes are still not fully understood. Therefore, water vapor measurements and retrievals are important for further understanding of the Earth's changing climate.

### 1.1.3 A Brief History of FIR Measurements and Atmospheric Retrievals

There exist several suggested methods for deducing atmospheric state variables such as water vapor content. Measurements of the infrared have been used for water vapor and temperature retrievals for years. The Nimbus satellite missions were key in demonstrating the concept of indirect measurements of temperature and water vapor from space, particularly the Satellite Infrared Spectrometer (SIRS), which was used for the first space-borne water vapor retrievals. The underlying principles behind such indirect measurements, or retrievals, is that the varying strength of water vapor absorption at different IR wavelengths changes the altitude from which the satellite receives emitted energy from the atmosphere. Thus by measuring emission in weak, moderate, and strong absorption lines, the satellite gathers information about the vertical temperature and water vapor profile.

While the far infrared may offer potential for improving water vapor retrievals, spectral measurements of the FIR have not been routine in the past. The last spectral measurements of the FIR taken from space were by the Russian Meteor-28 and Meteor-29 satellites launched in the 1970s (Spankuch and Dohler,1985; Timofeev et al., 2019). Since then, several instruments have measured FIR energy from space, but the measurements have been broadband rather than spectral. Unlike narrow band channels, broadband channels integrate information over a wide range of wavelengths, which are not conducive to retrieving the atmospheric state, as they mask the important variations in emission height needed for temperature and humidity retrievals. These measurements are commonly used in studying the Earth's radiation budget.

To date, the only routine spectral measurements of the thermal infrared have been from instruments such as the High resolution Infrared Radiation Sounder (HIRS), the Atmospheric Infrared Sounder (AIRS), the Cross-track Infrared Sounder (CrIS), the Infrared Atmospheric Sounding Interferometer (IASI), the Moderate Resolution Imaging Spectroradiometer (MODIS),

and others, all of which measure only in the mid infrared. These instruments have been regularly used for atmospheric retrievals. Spectral measurements of the FIR were limited to airborne and ground campaigns. Instruments of note include the Tropospheric Airborne Fourier Transform Spectrometer (TAFTS) which measured radiation between 80-800 $cm^{-1}$ (12-120 μm) with a spectral resolution of 0.1 $cm^{-1}$. TAFTS was designed for the purpose of taking direct observations of upper tropospheric humidity to understand the radiative properties of the UTH and cirrus clouds (Canas, Murray, & Harries, 1997; Warwick et al., 2022). Another airborne instrument was the Far Infrared Spectroscopy of the Troposphere (FIRST). This instrument measured between 50 and 2000 $cm^{-1}$, and the wider spectral range allowed for overlap with atmospheric sounders like AIRS, CrIS, and IASI, that measure the MIR (Mlynczak et al., 2002). Yet another instrument worth mentioning is REFIR or Radiation Explorer in the Far InfraRed. The REFIR instrument is a Fourier transform spectrometer measuring between 100 and 1100 $cm^{-1}$ with a resolution of 0.5 $cm^{-1}$. It was first deployed in a field campaign in 2004 (Palchetti et al., 2004). The REFIR instrument is the basis for the FORUM instrument currently in development (Palchetti et al., 2020).

There have been some ground-based measurements of the FIR, notably the Atmospheric Emitted Radiance Interferometer (AERI) located on several Atmospheric Radiation Measurement (ARM) sites (Ackerman & Stokes, 2003).

All of these instruments were used during several campaigns which demonstrated the importance of FIR radiation. For example, the Radiative Heating in Underexplored Bands Campaign (RHUBC) consisted of two campaigns measuring FIR radiation in regions of low atmospheric moisture to improve understanding of radiative processes in the upper troposphere and lower stratosphere and to refine water vapor spectroscopic models. The first campaign took place in the ARM site in Barrow, Alaska, and the second was in the Atacama Desert (Turner &

Mlawer, 2010; Mlawer et al., 2019). These campaigns, as well as other research conducted in the last several decades has highlighted the need for regular spectral measurements of FIR radiation from space.

## 1.2 The Need for FIR Measurements

One of the reasons for regular measurements is that the FIR plays an important role in climate. Radiation budgets and theoretical studies have shown that the FIR makes up a large portion of the radiation emitted to space from the Earth. In fact, globally, 45% of emitted radiation from Earth is found in the FIR. In polar regions, this proportion is even larger, reaching ⅔ over the coldest regions in Antarctica (Harries et al., 2008).

With regards to atmospheric retrievals, the FIR, or more specifically, the water vapor rotational absorption band found within the FIR portion of the spectrum, has stronger continuum and line absorption than the vibrational absorption band found in the MIR. As such, measurements of FIR are ideal for retrieval of atmospheric water vapor in areas of low water vapor amount, such as the mid to upper troposphere, stratosphere, and high latitudes near the surface. A comparison of information content between the FIR and MIR shows that when both spectral regions are measured with instruments having equivalent noise levels, the FIR measurement has an advantage in information content and vertical resolution for water vapor retrievals, and for temperature retrievals outside of the tropics (Merrelli and Turner, 2012). The reduced temperature information in the tropics in the FIR measurement is due to the lower troposphere becoming opaque at FIR wavelengths in atmospheres with large total water vapor amounts. Rizzi, Serio, and Morati (2002) had similar findings, in that the FIR enhanced the ability to retrieve water vapor throughout the mid to upper atmosphere and into the boundary layer in polar regions, but that the retrieval of lower troposphere water vapor was better in the

MIR in the tropics because of the large amounts of water vapor in the atmosphere. Shahabadi and Huang (2014) likewise found that the inclusion of FIR wavelengths was important for accurate measurements of stratospheric water vapor, and that in their study, uncertainties in other species of atmospheric gases such as $CH_4$, $N_2O$, $O_3$, and $CO_2$ did not significantly impact the water vapor retrieval, suggesting that FIR water vapor retrievals are less affected by interference from other trace gases.

## 1.3 Upcoming Missions

With the renewed interest in the FIR and technological advances in remote sensing technology, several satellites will be launched over the next several years that will take spectral measurements of the FIR. The Far Infrared Outgoing Radiation Understanding and Monitoring (FORUM) is a mission by the European Space Agency. The mission goal is to measure the FIR with high accuracy and spectral resolution. The satellite is currently slated for launch in 2026 (Palchetti et al., 2020). Another mission, from the Canadian Space Agency, is the Thin Ice Clouds in the Far InfraRed Experiment (TICFIRE). As its name suggests, TICFIRE will observe optically thin ice clouds in the Arctic using FIR measurements (Libois et al., 2015).

## 1.4 PREFIRE

Before either FORUM or TICFIRE, the Polar Radiant Energy in the Far InfraRed Experiment (PREFIRE) is scheduled to launch in 2023. PREFIRE will involve two 6U CubeSats in separate polar orbits. This will allow for overlap in spectral measurements at orbital intersections. The goal of PREFIRE is to study the FIR and Arctic processes at subdaily to seasonal timescales, something difficult to do with airborne and ground-based measurements.

Aboard the CubeSats, the Thermal infrared spectrometer (TIRS) will measure radiation between 5 and 54 μm with a spectral resolution sampling of 0.86μm (L'Ecuyer et al, 2021). A sample TIRS spectrum corresponding to a clear-sky scene from Greenland is shown in Figure 1.1.

This paper discusses an atmospheric retrieval algorithm developed for the TIRS instrument and the PREFIRE mission. The algorithm is a clear sky retrieval algorithm that uses Principal Component Regression (PCR) to retrieve temperature and humidity information. We will describe the theoretical basis for the algorithm, present some synthetic retrieval results, explore ways to improve performance using simple pre-screening of the algorithm database, and discuss the relative roles of MIR and FIR channels in constraining atmospheric temperature and humidity profiles.

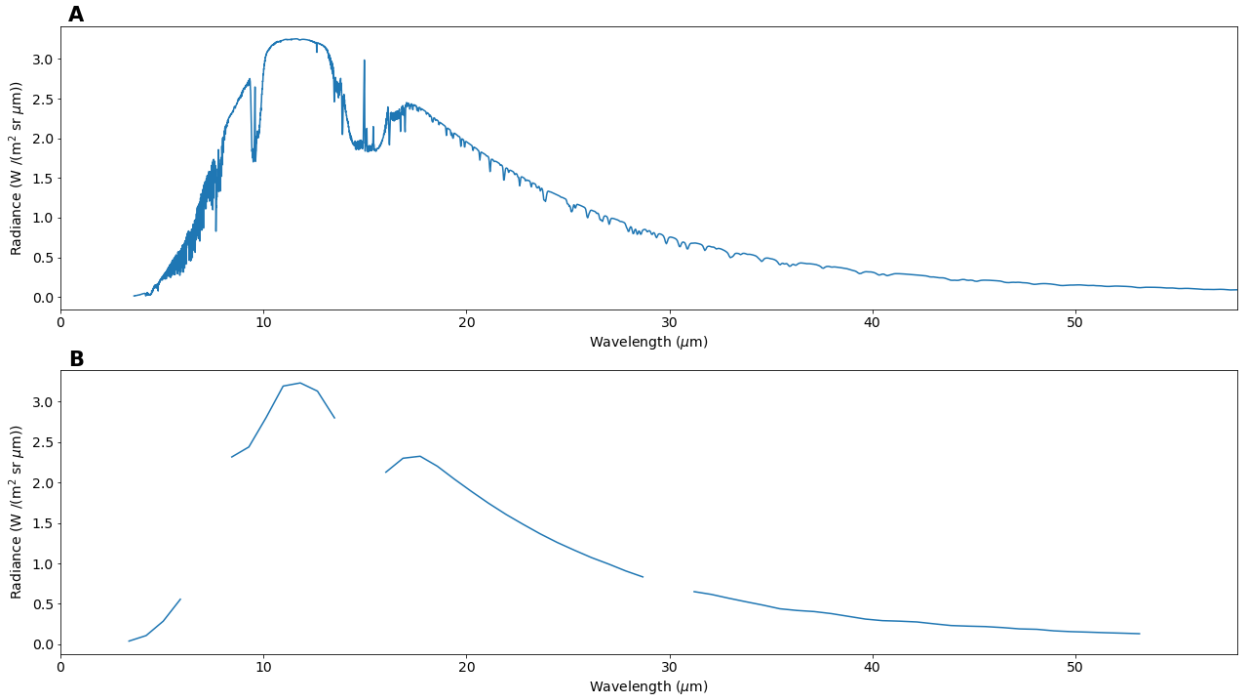


*Figure 1.1 – A simulated clear-sky radiance in high spectral resolution (A) and simulated from the TIRS instrument (B). The sample profile was taken over Greenland in March.*

# Chapter 2: Methodology

## 2.1 Retrieval Method

### 2.1.1 Principal Component Regression

The algorithm described here is based on a similar algorithm developed for the Atmospheric InfraRed Sounder (AIRS) as part of the International MODIS (Moderate Resolution Imaging Spectroradiometer) and AIRS Processing Package (IMAPP) (Weisz et al., 2007). Both algorithms retrieve atmospheric state variables by means of Principal Component Regression (PCR). The primary benefit of this retrieval method is speed and computational efficiency. Unlike other atmospheric retrieval methods, such as optimal estimation, PCR requires no first guess of the atmospheric state. In its most basic form, a retrieval algorithm using PCR only requires observations, with no additional information to produce a prediction of the atmospheric state with reasonable accuracy and precision, provided a large and representative sample of data is used to train the algorithm.

Radiative transfer equations used to relate atmospheric state variables to radiances are non-linear, so solving these equations for state variables is complex. PCR simplifies the problem by assuming a linear relationship between the predictors and predictands. The retrieval process is expressed by the following equations:

$$x_{ret} = CA^T + \bar{x}_{train}$$

$$A = (y_{obs} - \bar{y}_{train})U$$

Vector $x_{ret}$ is the retrieved state vector, also known as the predictand. $C$ is the matrix of regression coefficients, $A$ is the matrix of principal components (PCs), and $x_{train}$ is the average state vector of the training data. The principal components, $A$, are calculated by subtracting the average $y$ vector from the observations and applying matrix $U$, which is the truncated

eigenvector matrix. The means by which these quantities are obtained are described in section 2.2.

### 2.1.2 Principal Components

The shape of matrix A in the retrieval equations is [1 x *npc*], where *npc* is the number of principal components retained for the retrieval. As shown in the retrieval equations, the principal components are determined by applying a truncated eigenvector matrix *U* to the residuals of the observations from the average state. The maximum number of principal components one may use is equal to the number of predictors. For a basic PCR retrieval algorithm, this means the number of channels measured by the instrument. Often, using the maximum number of principal components is unnecessary, and in some cases may be ill-advised. As is demonstrated by Weisz et al. (2007), use of too many principal components can result in over-fitting of the retrieved variables, increasing retrieval error.

There are several methods for determining the number of principal components to retain. One method is to examine the eigenvalues for the calculated eigenvectors. The eigenvalues give the relative weight of importance for each eigenvector, as they are proportional to the variance in the data explained by their respective eigenvector. One may determine the number of principal components to retain based on the eigenvalues and selecting those which explain a certain amount of variance.

The number of PCs is more important in retrievals applied to noise-free radiance measurements than in retrievals applied to noisy measurements. As mentioned previously, retaining too many PCs can result in increased retrieval error due to overfitting. In a retrieval applied to noisy measurements, the additional error created by adding additional PCs beyond the

optimal number is less than it is for noise-free measurements, because the additional information is partially masked by the added noise.

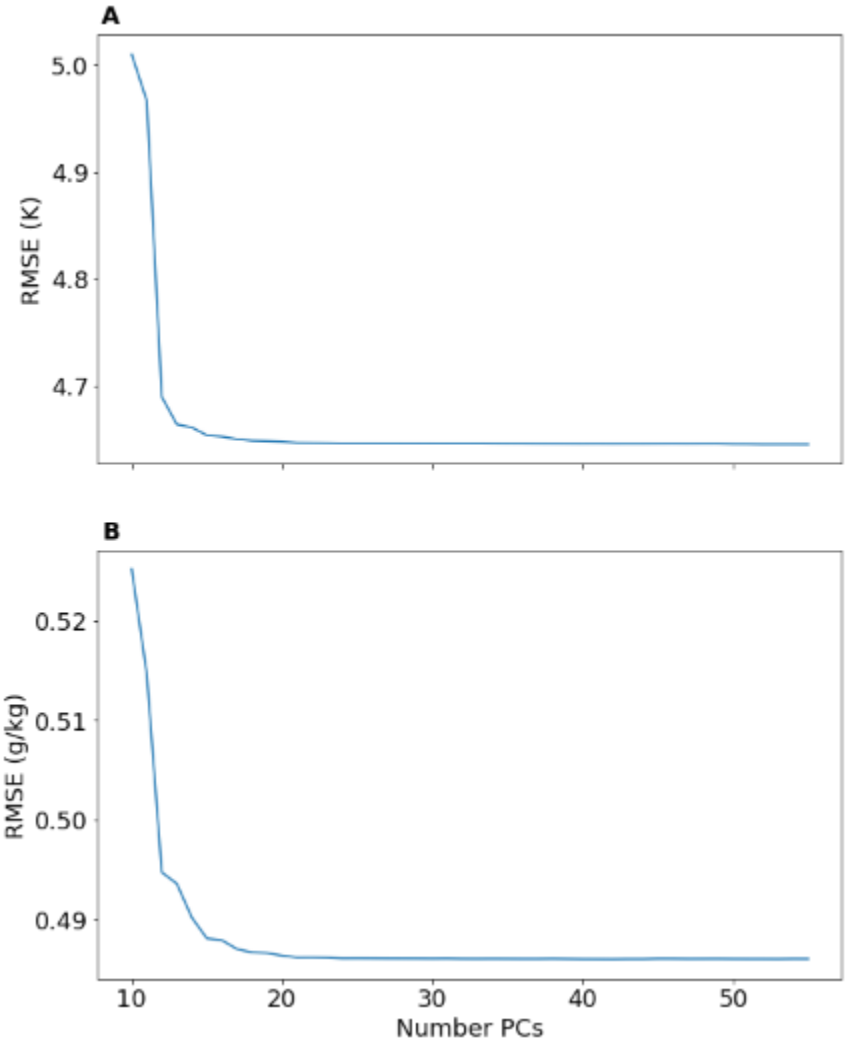


*Figure 2.1 – Column-averaged RMSE for simulated temperature (A) and mixing ratio (B) retrievals from the TIRS instrument, as a function of number of principal components retained.*

An alternative method, and the method used for this case, is to select the number of PCs based on the vertically averaged root mean square error (RMSE) between calculated atmospheric state and the true values. RMSE is the standard deviation of the retrieval error, if the retrieval bias is equal to zero. In other words, it is a measure of retrieval precision. By choosing the number of principal components that minimizes the RMSE, we are maximizing the precision of

the retrieval. Figure 2.1 shows the vertically averaged temperature and water vapor RMSE from a TIRS retrieval as a function of number of PCs. The average RMSE drops quickly as number of principal components increases, and then becomes approximately constant. Indeed, the average RMSE becomes constant out to two decimal places beyond 15 PCs for temperature and beyond 19 PCs for water vapor. Therefore, to minimize RMSE while maximizing computational efficiency, 20 PCs were selected for this retrieval.

It should be noted that in using this method, we are assuming that the retrieval is accurate, or that the retrieval bias is equal to zero. If this were not the case, reducing the RMSE could increase the retrieval error by reducing the spread of errors around an inaccurate average error. However, it will be shown later in this paper that the assumption of a bias of zero is reasonable.

### 2.1.3 Predictors and Predictands

Along with radiance, the retrieval also uses surface pressure as a predictor. Surface pressure is included as a predictor to aid in retrieval performance across regions with different surface altitudes, especially in the large Greenland and Antarctic ice sheets. For the simulated retrievals presented in this paper, surface pressure is given in the training dataset, while radiance is calculated using a radiative transfer model. In operation, the measured radiances will be used, and surface pressure will be provided by auxiliary data from interpolated numerical weather model analysis fields.

The variables retrieved by the algorithm are skin temperature, total precipitable water, water vapor mixing ratio profile, and temperature profile. Profile data is retrieved at 101 pressure levels between 1100 and 0.005 hPa.

## 2.2 Algorithm Training

As previously mentioned, PCR simplifies the atmospheric retrieval problem by assuming a linear relationship between the predictors and predictands:

$$X = CY^T$$

The best-fit solution based on the least squares method is one that minimizes the sum of the squared residuals:

$$\sum(X - CY^T)^2$$

Through minimization, we determine an expression for the regression coefficients:

$$C = XY(Y^TY)^{-1}$$

In PCR, rather than directly using the observations (*Y*), regression coefficients are calculated from principal components.

The training and calculations are performed on the residuals of the training data:

$$X' = X_{train} - \bar{x}_{train}$$

$$Y' = Y_{train} - \bar{y}_{train}$$

$X_{train}$ is a matrix of the state vectors for the training data, and $Y_{train}$ is a matrix of the corresponding predictors (radiance and surface pressure). *X'* and *Y'* represent the residuals, calculated by subtracting the average of the training data from the training data itself.

The eigenvectors are calculated from the covariance of the predictors. The eigenvectors are then ordered by decreasing eigenvalue and truncated to the number of PCs retained. At this point, the eigenvector matrix has dimensions of [*ny* x *npc*], where *ny* is the number of predictands, and *npc* is the number of principal components.

It is after this step that noise is added to the residuals of the radiance data to account for noise in observations. Figure 2.2 shows the noise equivalent delta radiance (NEDR) for the TIRS

instrument. Noise is added to the radiances randomly, assuming a normal distribution with an average of zero and a standard deviation equal to the NEDR at a given TIRS spectral channel.

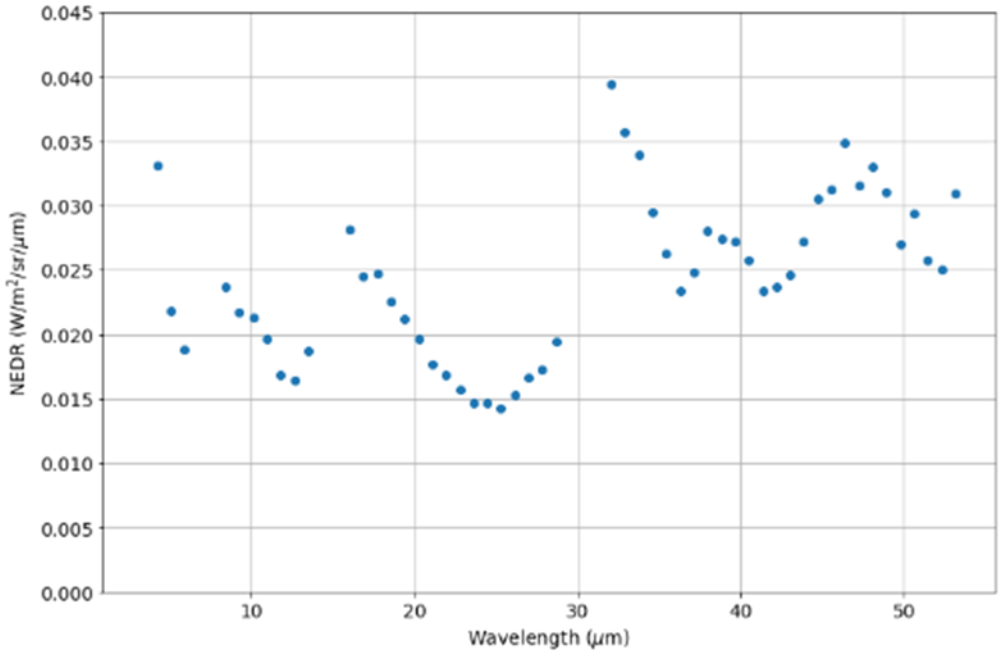


*Figure 2.2 – Estimated noise equivalent delta radiance (NEDR) in W $m^{-2}$ $sr^{-1}$ $\mu m^{-1}$ for each TIRS channel*

Finally, the regression coefficients are calculated:

$$C = X'^{T} A (A^{T} A)^{-1}$$

$A$ is the matrix of principal components, calculated as:

$$A = Y'U$$

where $U$ is the truncated eigenvector matrix.

## 2.3 Data Sources

### 2.3.1 Training Dataset

The atmospheric data used for training and testing the algorithm comes from the SeeBor database. This database contains global atmospheric profile data from 1968-2005. Data sources

include NOAA-88b, Thermodynamic Initial Guess Retrieval - 3 (TIGR-3), radiosondes, ozonesondes, and the European Center for Medium-Range Weather Forecasts (ECMWF). Each profile includes measurements of temperature, water vapor mixing ratio, surface pressure, skin temperature, total precipitable water, wind, and concentrations of ozone (O3), carbon monoxide (CO), methane (CH4), sulfur dioxide (SO2), and nitrous oxide (N2O). The database also contains latitude and longitude for each profile, a land/water flag, International Geosphere-Biosphere Programme (IGBP) land cover classification, and surface emissivity values at 10 key hinge points (Borbas et al., 2005). These hinge points come from the University of Wisconsin (UW) Baseline Fit Emissivity Database and can be used to infer high spectral resolution surface emissivity. Section 2.4.2 discusses emissivity data in further detail. All profile data from the database contains values at 101 pressure levels between 1100 and 0.005 hPa, the same pressure levels used in the retrieval, so no interpolation was required. At pressure levels greater than the surface pressure, the profile data is given the same value as the surface data. For sounding data that does not contain measurements up to the top of the atmosphere, standard atmosphere data was used (Borbas et al., 2005). For the PCR retrieval, the forward model calculations are run using a constant carbon dioxide (CO2) concentration of 400ppm for all data. This number was chosen to be more representative of a modern value.

A total of 15,704 clear-sky profiles were used for training and testing the retrieval. One tenth, or 1,571, of the profiles are set aside for testing, while the remaining 14,133 were used for training. For brightness temperature classification (BTC, see section 2.5), the same test and training set were used as in the base retrieval, but the datasets were categorized into different classes according to their 11 μm brightness temperature. The number of profiles for training and testing each class in BTC are outlined in Table 2.2. The numbers displayed in the table point out a bias in the training data. Though SeeBor is a global dataset, the warmest classes contain more

profiles than the colder classes. This is due in part to the type of data used. By using ground-based sounding data, the database is constrained by data availability to the regions where soundings are typically taken. This means, there are more soundings in more heavily populated regions, such as the midlatitudes and tropics, as shown in Figure 2.3. It will be shown in section 3.1 that the retrieval still performs within expected limits, even in cold regions.

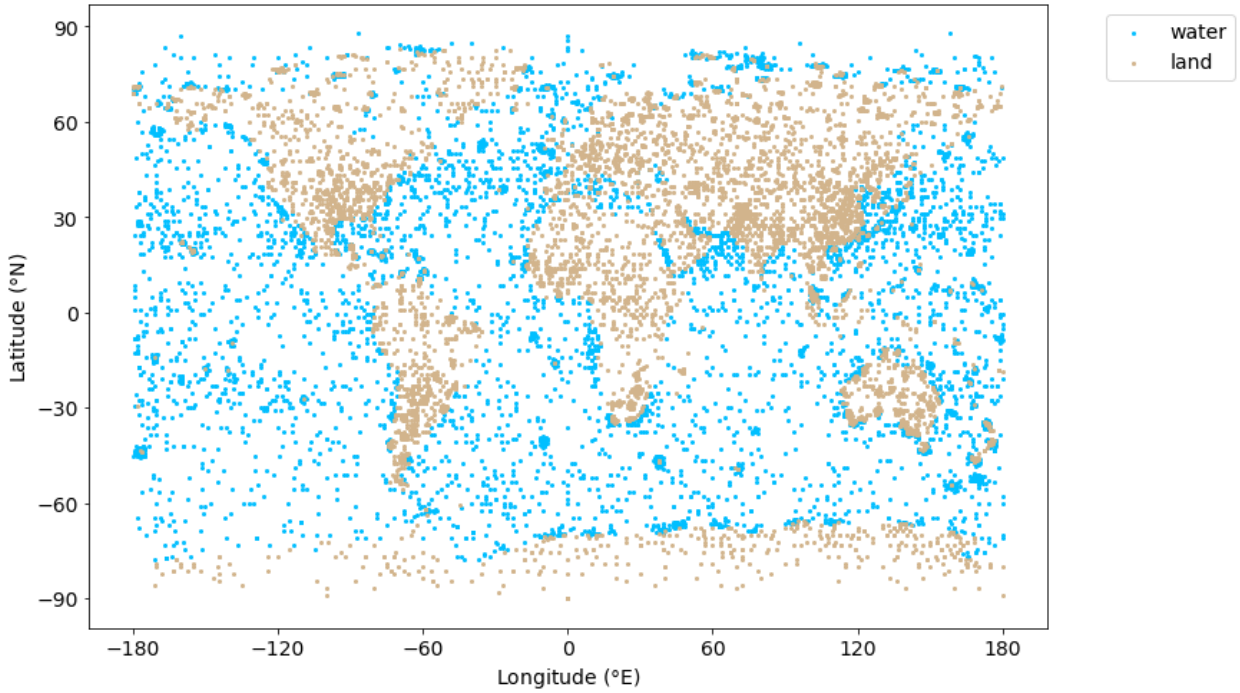


*Figure 2.3 – Locations of profiles from the SeeBor database.*

### 2.3.2 Emissivity Data

To simulate radiances corresponding to the training data, a forward model was run using the data from the SeeBor database. The forward model also required surface emissivity data. To fulfill this requirement, we designed an emissivity library based on land cover type. The sources of emissivity data for the library, and the method of creation are described in this section.

#### *2.3.2.1 Land Emissivity*

As mentioned previously, the SeeBor database contains surface emissivity at 10 wavelengths (3.6, 4.3, 5.0, 5.8, 7.6, 8.3, 9.3, 10.8, 12.1, and 14.3 microns) for each profile. These hinge points come from the University of Wisconsin (UW)/Cooperative Institute for Meteorological Satellite Studies (CIMSS) Baseline Fit Emissivity database (Seeman et al., 2007). The corresponding UW InfraRed Emissivity (UWIREMIS) algorithm for deriving high spectral resolution emissivity from these hinge points was used for surface emissivity data over land. A separate method was used to determine emissivity over water.

The UWIREMIS algorithm calculates high spectral emissivity from the hinge points using principal component regression. Regression coefficients are calculated from laboratory emissivity measurements and applied to the hinge point emissivities to get emissivities at 416 wavelengths between 3.6 and 14.3 microns. From there, the emissivity is interpolated to the wavelengths required by the forward model using linear interpolation (Borbas et al., 2007).

#### *2.3.2.2 Ocean Emissivity*

Though UWIREMIS was used for land data, the SeeBor database included the emissivity data for all profiles flagged as being over water. As such, a separate model was used to determine emissivity over ocean. The emissivity model is a pre-calculated table based on wind speed and sensor zenith angle, with emissivity values between 850 and 2700 $cm^{-1}$, or 3.7 and 11.8 μm. Wind speed is required to account for surface roughness, and sensor zenith angle and the complex index of refraction are used to approximate effective incidence angle of radiation (Nalli, Minnet, and Van Delst, 2008). The wind speeds used for the model are those provided in the SeeBor database, and the sensor zenith is held at 0° for all profiles. As with the UWIREMIS

algorithm, the returned emissivities are interpolated to the wavelengths needed for the forward model with linear interpolation.

#### *2.3.2.3 Far Infrared Emissivity*

Both the land and ocean emissivity algorithms provide emissivity data only in the mid infrared. UWIREMIS provides emissivities between 3.6 and 14.6 μm, and the Nalli, Mannet and Van Delst algorithm contains emissivity values between 3.7 and 11.8 μm. Therefore, a separate model was required to include emissivity in the far infrared (Huang et al, 2016). This model, which will be referred to as H16, provides emissivity values between 10 and 2000 $cm^{-1}$ (5 – 1000 μm). H16 uses first-principle calculations to determine emissivity for water, ice, snow (fine, medium, and coarse), and desert. The desert class contains 16 subclasses with variations in the effective radii of silt between 30 and 45 μm, accounting for differences in sand particles. For the sake of simplicity, the subclass with 30 μm effective radius is used.

Due to the complexities of vegetation surfaces, first-principle calculations were not used for vegetation surfaces. Instead, measured emissivities from the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) database were used to define four different vegetation classes: grass, dry grass, deciduous, and conifer. Since ASTER does not measure the far infrared, far infrared wavelengths are assigned a constant emissivity value. An additional class is defined as a combination of grass and desert, for those land surfaces in which desert regions are covered by vegetation. This emissivity is calculated as 45% grassland (meaning grass and dry grass), and 55% desert, using the average emissivity of all desert subclasses (Huang et al., 2016).

#### *2.3.2.4 Creation of the Emissivity Library*

In the initial stages of algorithm development, FIR emissivity was held constant for all profiles. Once the algorithm was fully functional, FIR emissivities were included in the algorithm training process. This was done by merging the H16 emissivities with the UWIREMES and Nalli, Minnet and van Delst data sets. Rather than simply appending the FIR emissivity values from H16 to the other MIR emissivities, the datasets were spliced together at 1750 $cm^{-1}$ or 5.7 μm, which is near the center of the MIR water vapor absorption band, and would therefore not be sensitive to any sharp discontinuities created by the splice. Figures 2.4 and 2.5 show examples of the spliced emissivities for grasslands and snow respectively.

As all the SeeBor profiles were assigned an IGBP class, which contains 19 different surface classifications, the emissivity library used for the forward model contains classes based on the IGBP classifications. However, additional classes needed to be defined to match up the mid infrared emissivities based on the 19 different IGBP classes with the 11 classes defined for the FIR emissivity data. The classes used in the emissivity library as well as the SeeBor IGBP classes and H16 land types used are listed in table 2.1.

| New Class | IGBP Classes | H16 Classes |
|---|---|---|
| **evergreen needleleaf** | evergreen needleleaf | Conifer |
| **evergreen broadleaf** | evergreen broadleaf | Deciduous |
| **deciduous needleleaf** | deciduous needleleaf | Conifer |
| **deciduous broadleaf** | deciduous broadleaf | Deciduous |
| **mixed forest** | mixed forest | Deciduous |
| **closed shrublands** | closed shrublands | Grass |
| **open shrublands** | open shrublands | grass<br>dry grass<br>desert-silt re= 30 micron |
| **woody savannas** | woody savannas | grass<br>dry grass |

| | | |
|---|---|---|
| **Savannas** | savannas | grass<br>dry grass |
| **grasslands** | grasslands | grass<br>dry grass |
| **permanent wetlands** | permanent wetlands | Grass |
| **croplands** | croplands | grass<br>dry grass |
| **urban and built-up** | urban and built-up | grass<br>dry grass |
| **cropland mosaics** | cropland mosaics | grass<br>dry grass |
| **fine snow** | snow and ice | fine snow |
| **medium snow** | snow and ice | medium snow |
| **coarse snow** | snow and ice | coarse snow |
| **bare soil and rocks** | bare soil and rocks | grass<br>dry grass<br>desert-silt re= 30 micron |
| **water bodies** | water bodies<br>(land flag = 0) | pure water |
| **Coastline** | water bodies<br>(land flag = 1) | grass<br>dry grass |
| **Tundra** | tundra | grass<br>dry grass |

*Table 2.1 – Table of classes defined for the emissivity library created for the retrieval*

An additional class was added into the library for profiles along shorelines. In the SeeBor dataset, 739 data points showed conflicts between the land flag and IGBP class. A profile over water should be given an IGBP class of 17, and additionally should have a land flag of 0, indicating no land surface. However, 646 profiles flagged as land had an IGBP classification of 17, for water bodies. Furthermore, the hinge point emissivities for these profiles were inconsistent with typical emissivity values for water surfaces, many of them showing characteristics of grass or other land types. As such, these profiles were assumed to be mis-classified land profiles. For these profiles, and additional “coastline” class was created. An

additional 93 data points were flagged as water, but were given an IGBP class for other land cover types. In this case, these profiles were determined to be mis-classified water profiles. The emissivities of these points were consistent with water and were therefore assumed to be profiles over water bodies.

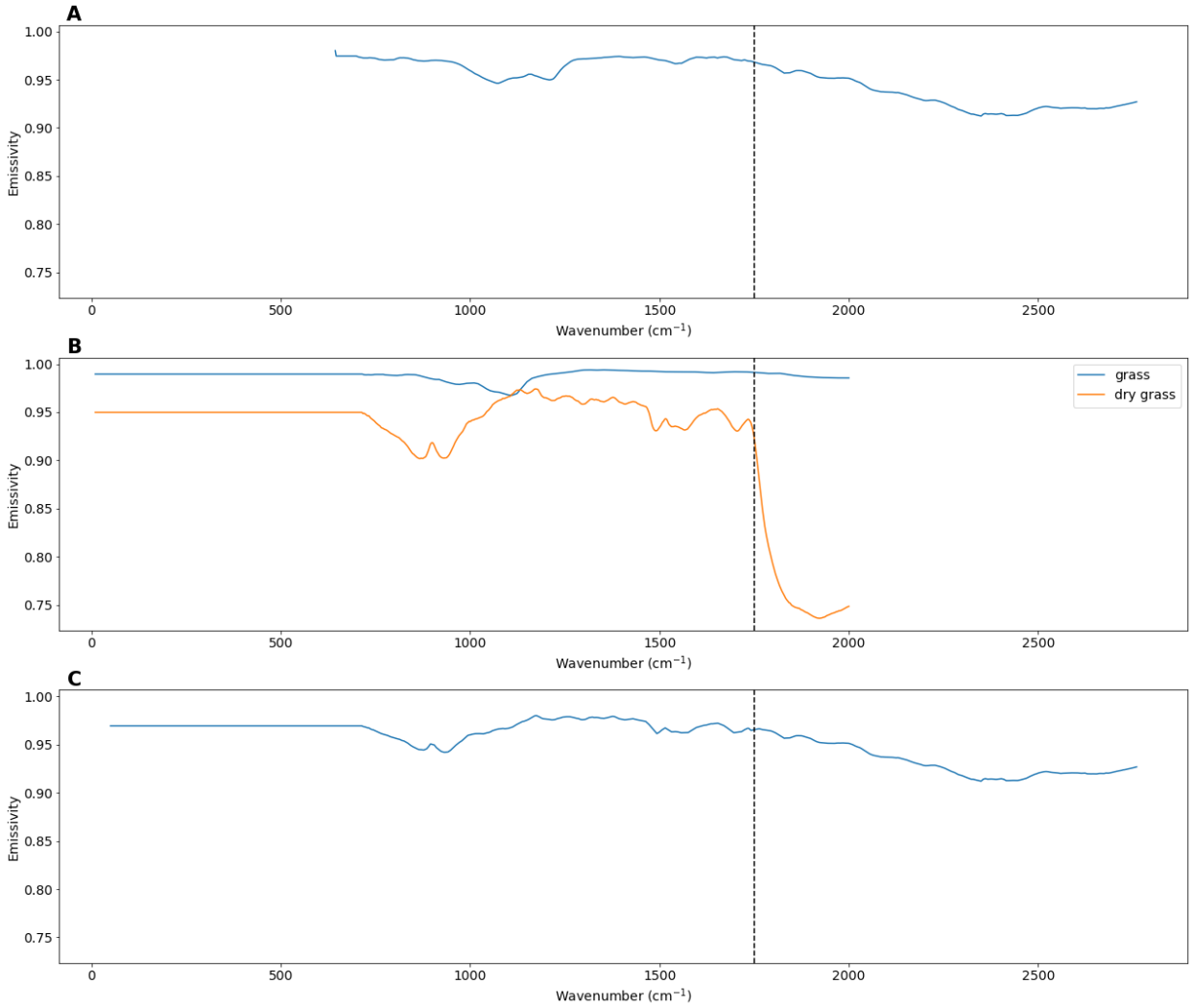


*Figure 2.4 – Grassland emissivity from the SeeBor dataset (A), grass and dry grass emissivity from the H16 emissivity dataset (B), and the combined emissivity used in the new emissivity library. The dashed line indicates the wavenumber at which the datasets were spliced together. The average of the grass and dry grass emissivities from the H16 dataset are used to create the new emissivity.*

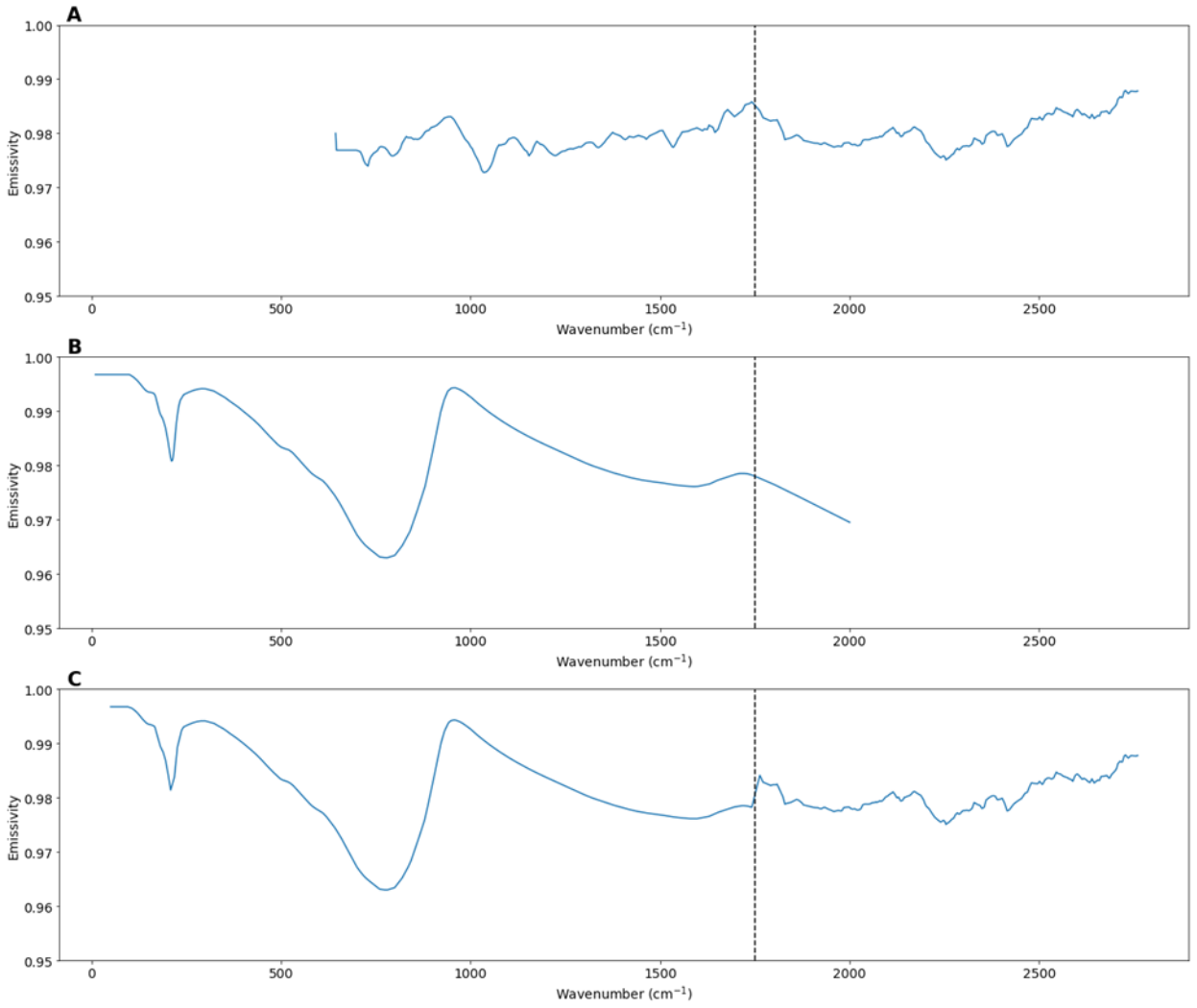


*Figure 2.5 – Snow and ice emissivity from the SeeBor dataset (A), medium snow emissivity from the H16 dataset (B), and the medium snow emissivity used in the retrieval, obtained by combining the other two emissivities (C). The dashed line indicates the wavenumber where the datasets are spliced together.*

## 2.4 Radiance Data and Forward Model

To train and test the algorithm, radiance data corresponding to the SeeBor training data was required. The forward model used was the Principal Component Radiative Transfer Model (PCRTM), which as the name suggests, is a Principal Component based method to model instrument radiance measurements from environmental data (Liu et al, 2006).

As TIRS is a new instrument, there is no radiative transfer model to directly simulate TIRS measurements. To create TIRS simulated measurements, PCRTM was run to output radiances at the wavenumbers for the Climate Absolute Radiance and Refractivity Observatory (CLARREO) mission. The CLARREO radiances range between 50 and 2760 $cm^{-1}$ with a sampling of 0.5 $cm^{-1}$ spectral resolution of 1 $cm^{-1}$ (Wielicki et al., 2013). Spectral response functions (SRF) for the TIRS spectral channels were applied to the CLARREO radiances to simulate TIRS measurements. Figure 2.6 shows an example of this procedure from a random profile chosen from the SeeBor dataset.

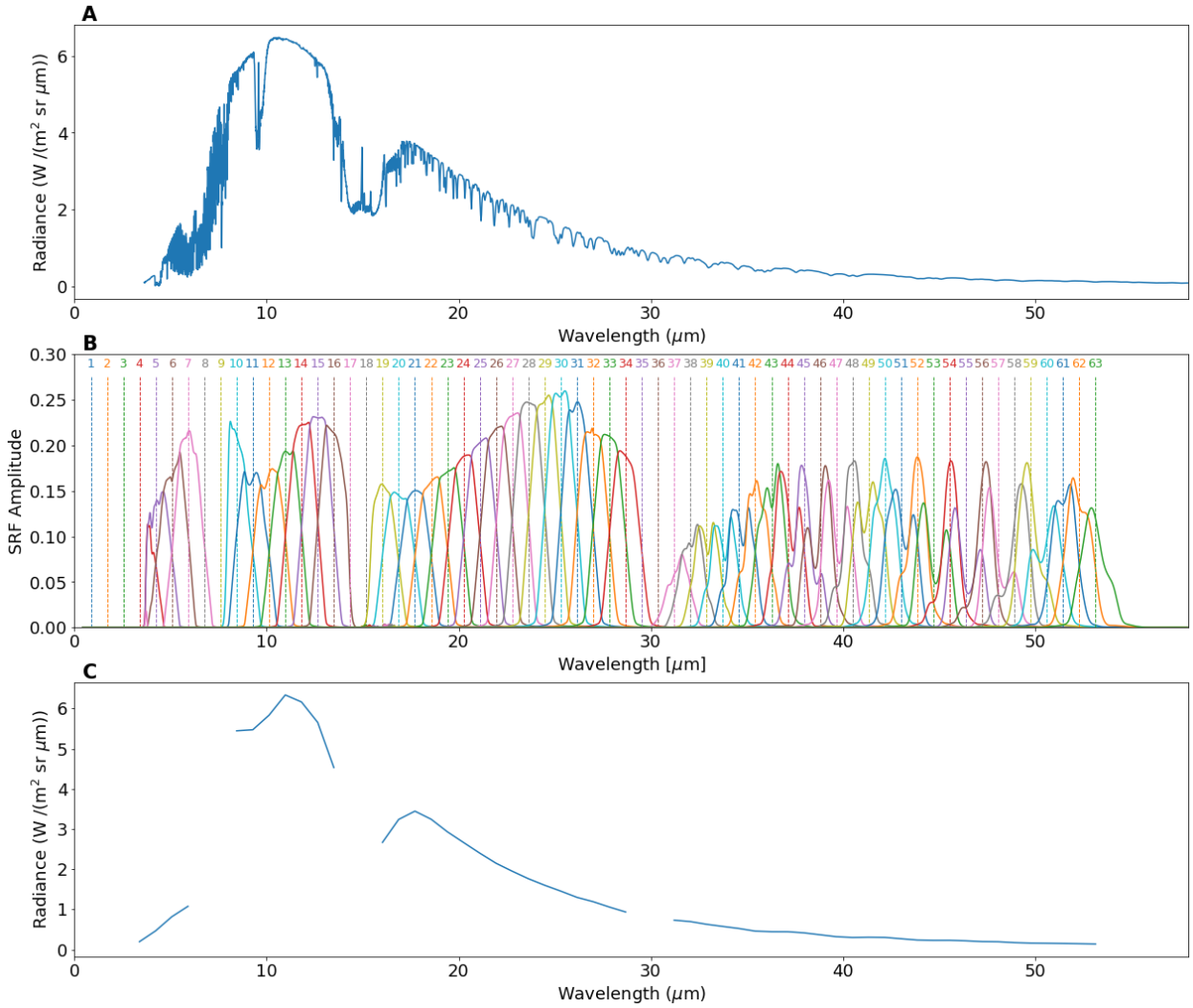


*Figure 2.6 – An example of CLARREO radiances for a randomly selected SeeBor profile calculated from the PCRTM forward model (A). (B) shows the TIRS spectral response function, and (C) shows the corresponding TIRS radiance when the SRF is applied to the CLARREO radiance*

## 2.5 Brightness Temperature Classification

To improve algorithm performance, we further classify profiles according to their MIR window channel brightness temperatures in the retrieval process. Brightness Temperature Classification (BTC) is designed to improve retrieval performance for different regions by selecting a set of regression coefficients and eigenvectors based on measured brightness temperature in a window channel or channels. Window channels are channels that measure radiance in a region of the electromagnetic spectrum where there is little to no atmospheric

absorption of outgoing radiation from the Earth's surface. As such, the measured brightness temperature from these channels is generally directly related to the surface temperature. BTC works by classifying measured radiances into different categories based on measured brightness temperature in the window channel.

High spectral resolution instruments such as AIRS, upon which this method is based, typically have many channels within an atmospheric window. For the AIRS algorithm, the brightness temperature classification is determined by taking the average of 11 channels, all located in a window region of the spectrum centered at approximately 11 μm. The advantage of using more than one channel is the reduction of the instrument noise and the effects of weak water vapor lines through averaging. TIRS, having a lower spectral resolution, does not have this advantage, and only has one channel located in the window region. This one channel, however, is located at 10.97μm, near the center of the window region. Figure 2.7 shows a plot of simulated brightness temperature measurements at the 10.97 μm channel from TIRS compared to skin temperatures from the training data. Strong agreement exists between the channel measured brightness temperature and skin temperature from the training dataset. Only at the warmest skin temperatures does the data spread out away from the 1-1 line. Here, the brightness temperature measurements underestimate the surface temperature due to weak absorption by water vapor continuum in the window channel. In colder regions, the atmosphere has less moisture content, so the amount of absorption is negligible. It is only in the warmest regions where the atmosphere has a higher amount of moisture that water vapor absorption has a large enough signal that measured brightness temperatures deviate from skin temperature. Variations in surface emissivity can also cause deviations from the 1-1 line. However, the range in emissivity in this spectral region is small, between 0.95 and 0.99, and therefore unlikely to create the strong deviations seen at the warmest skin temperatures.

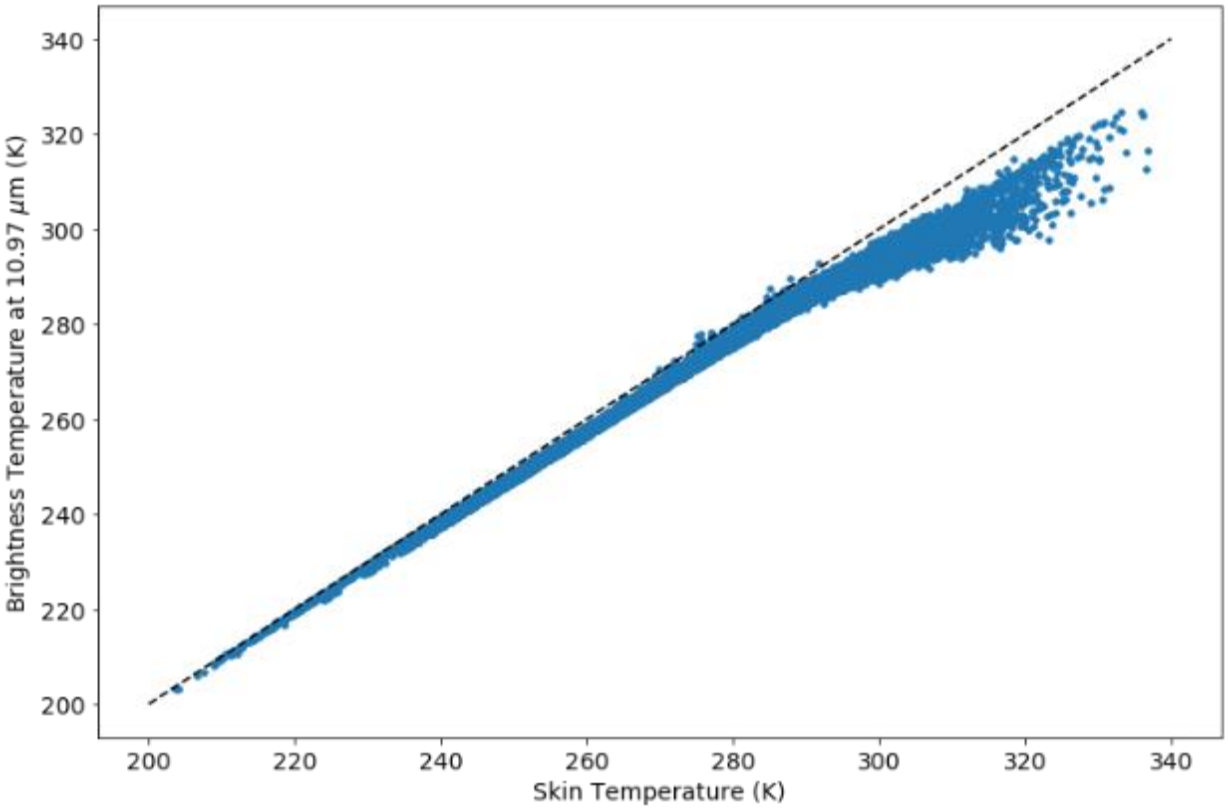


*Figure 2.7 – Simulated measured brightness temperature at the 10.97 TIRS channel compared to skin temperature reported in the SeeBor database. The dashed line represents a 1-1 correspondence.*

As for the process of classification, six classes are defined following the criteria defined by Weisz et al. (2007) in their IMAPP PCR retrieval. The brightness temperature thresholds for each class can be found in Table 2.2. There are six classes, the middle four of which span a range of 10K, and the first and last class acting as "catch-alls" for brightness temperatures lower or higher than the minimum and maximum temperatures used in the class definitions. In training the algorithm, a different range of brightness temperatures are used to classify a profile than is used in the retrieval. For training, the classes overlap with each other, as shown in Table 2.2a. This is done to minimize the error resulting from misclassification. The algorithm is trained the same way as in the base retrieval case, but a different set of regression coefficients, eigenvectors, and average states is calculated for each class. Then the retrieval is run by first classifying the

measurements into the classes defined in Table 2.2b and using the regression coefficients for that particular class.

**A**

| Class Number | Brightness Temperature Cutoffs (K) | Number of Profiles |
|---|---|---|
| 1 | BT ≤ 260 | 1966 |
| 2 | 250 < BT ≤ 270 | 1960 |
| 3 | 260 < BT ≤ 280 | 3294 |
| 4 | 270 < BT ≤ 290 | 5877 |
| 5 | 280 < BT ≤ 300 | 8189 |
| 6 | 290 < BT | 5051 |

| Class Number | Brightness Temperature Cutoffs (K) | Number of Profiles |
|---|---|---|
| 1 | BT ≤ 255 | 184 |
| 2 | 255 < BT ≤ 265 | 86 |
| 3 | 265 < BT ≤ 275 | 174 |
| 4 | 275 < BT ≤ 285 | 313 |
| 5 | 285 < BT ≤ 295 | 633 |
| 6 | 295 < BT | 181 |

*Table 2.2 – Brightness temperature class definitions for the training dataset (A) and test dataset (B)*

## 2.6 Analysis Methods

### 2.6.1 Comparison with Other Retrievals

As a means to qualitatively analyze the retrieval performance, the algorithm was also trained to run retrievals from the Infrared Atmospheric Sounding Interferometer (IASI) instrument. IASI is a Michelson interferometer first launched on the Meteorological Operation (MetOp)-A satellite by the European Organization for Exploitation of Meteorological Satellites (EUMETSAT). IASI is a high spectral resolution instrument, with measurements covering a spectral range of 645 – 2760 $cm^{-1}$ (3.62 – 15.5μm) with a spectral sampling of 0.25 $cm^{-1}$. IASI was designed for the purpose of high-resolution atmospheric soundings, and so provides a good means of comparison for TIRS retrieval performance (Hilton et al., 2012).

### 2.6.2 Error Measurement

The primary methods for analyzing retrieval performance are bias and standard deviation of error (SDE) The two quantities are calculated as follows:

$$Bias = \frac{\sum(x_{truth} - x_{ret})}{N}$$

$$SDE = \sqrt{\frac{\sum((x_{truth} - x_{ret}) - (\overline{x_{truth} - x_{ret}}))^2}{N - 1}}$$

Here, $x_{truth}$ is the state vector from the SeeBor dataset, $x_{ret}$ is the retrieved state vector, and $N$ is the number of data points, or test retrievals. Bias and SDE are calculated for each of the 101 retrieved pressure levels. In instances where we compare bias from different retrievals, the magnitude, or absolute value of bias is used. It is apparent from these formulas that bias represents the average retrieval error and SDE demonstrates the spread of retrieval error. As such, bias is a measure of retrieval accuracy, while SDE measures precision.

# Chapter 3: Results

## 3.1 Base Retrieval

In the initial base retrieval, we assess the performance of TIRS through measurements of retrieval bias and SDE, and by comparing the TIRS retrieval performance with that of IASI. By base retrieval, we refer to the most basic retrieval without brightness temperature classification, where the same set of retrieval coefficients are used for all retrievals. The retrieval results are compared to climatology, represented by the mean and standard deviation of all profiles in the SeeBor database. We shall discuss the retrieval performance of all four retrieved variables: temperature, water vapor mixing ratio, skin temperature, and total precipitable water.

Figure 3.1 shows the bias and error for the base retrieval when all testing profiles are used. Temperature retrieval is shown in figure 3.1a. The TIRS retrieval, indicated by the orange line, has a retrieval bias between –0.1 and 0.1 K throughout the troposphere. The largest bias occurs around 10 hPa, where the bias is around –0.3 K. Given that the retrievals will primarily focus on the troposphere and lower stratosphere, this peak in bias is not a cause for concern. Moreover, given that the bias is less than 1 K, the retrieval performance is reasonable, so we may consider that the temperature retrieval is accurate. To fully understand retrieval performance, we also compare the TIRS retrieval to the bias for IASI and the average SeeBor profile from the training dataset, also shown in figure 3.1. This average profile (green) is the average profile from the training data set representing the scenario where no measurement is made to constrain the estimate. If the PREFIRE retrieval has no skill in retrieving the temperature and humidity profile, it would return the average of the training profiles. In the pressure levels of interest, we can see that the retrieval bias is small for all three retrievals, on an order of magnitude of $10^{-1}$ K. As such, focus on SDE as a measure of retrieval performance and precision.

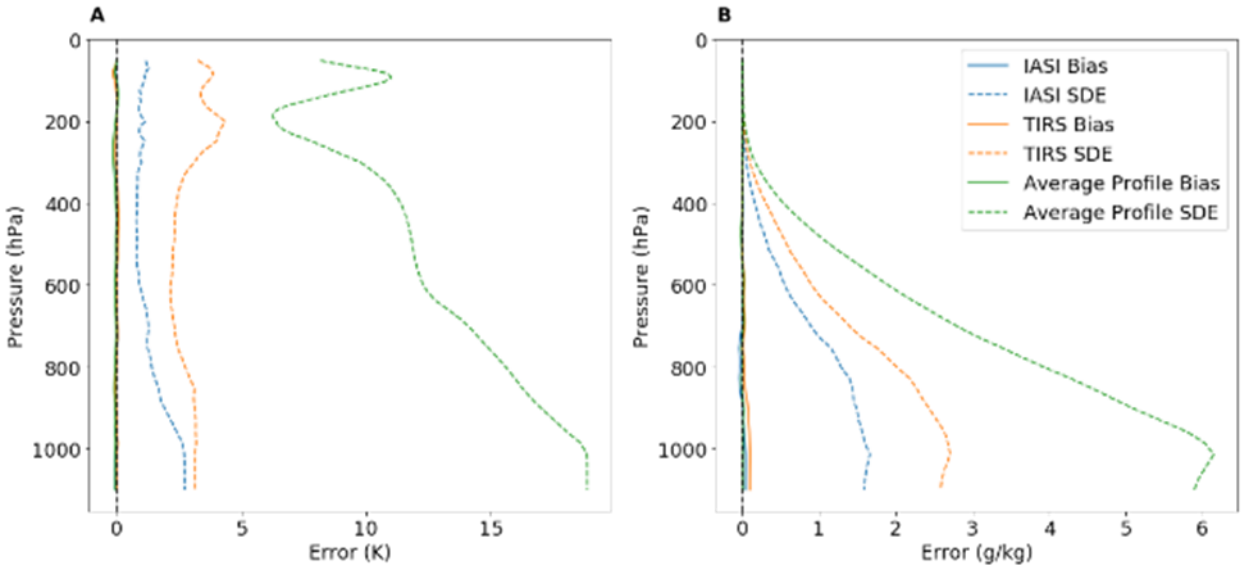


*Figure 3.1 – Base retrieval error for temperature retrieval (A) and mixing ratio retrieval (B) for all profiles in the validation dataset. Solid lines show retrieval bias, while dashed lines indicate standard deviation of error (SDE).*

The dashed lines in figure 3.1a show the SDE for temperature retrievals. Within the troposphere and lower stratosphere (up to about 40 hPa), the SDE for the TIRS retrieval is within 5 K. The SDE for IASI is within about 3 K, while the benchmark or Average Profile SDE is greater than 5 K, between about 6 and 18 K, with the highest SDE near the surface. This is to be expected, as global temperature variations are highest at the surface. Given that the TIRS SDE is more similar to that of IASI than it is to the benchmark, despite the lower spectral resolution, we can say qualitatively that the TIRS temperature retrieval is both reasonably precise and accurate, or that it is performing within expected variability.

For water vapor mixing ratio, we will see similar results to that of temperature. The mixing ratio retrieval biases are shown in Figure 3.1b. The TIRS retrieval bias peaks at high pressures near the surface, with a maximum magnitude of around 0.1 g/kg. Retrieval bias decreases going up in the atmosphere, due to the lack of moisture at higher levels. As to be expected, given its high spectral resolution, IASI has lower magnitude biases than TIRS. At pressures above 600 hPa (closer to the surface), the benchmark average profile retrieval has a

lower magnitude bias compared to both TIRS and IASI. However, the magnitude of the biases in both the TIRS and IASI retrievals is sufficiently small to give confidence that the retrievals are largely unbiased. Thus, it is not the accuracy of the retrieval, but the precision that needs to be analyzed. To do this, we consider the retrieval SDE, shown by the dashed lines in figure 3.1b. The SDE curve for TIRS has a similar shape to that of the bias curve: the peak SDE is near the surface, and SDE decreases with decreasing pressure. Moreover, the SDE is an order of magnitude higher than the bias, meaning that even though the retrieval tends to overestimate water vapor mixing ratio near the surface, the true value is still within the expected range of errors for the retrieval. The same may be said for IASI. As such, we may again conclude that the retrieval is reasonably accurate. As far as the precision of the TIRS retrieval, like with the temperature retrieval, the mixing ratio SDE for TIRS is between that of IASI and the average profile. So as with the temperature retrieval, the TIRS mixing ratio retrieval is performing within expectations.

The final two retrieved variables are skin temperature and total precipitable water. The bias and SDE values for these retrieved variables are shown in Figure 3.2. Once again, the biases in both retrievals are small, and consistent with zero bias. Small variations in the retrieval biases may be attributed to sampling noise. As with other retrieved variables, the TIRS SDE was between that of IASI and the average profile. For total precipitable water (TPW), the retrieval performance was similar to that for water vapor mixing ratio, perhaps to be expected as they are both measurements of atmospheric moisture.

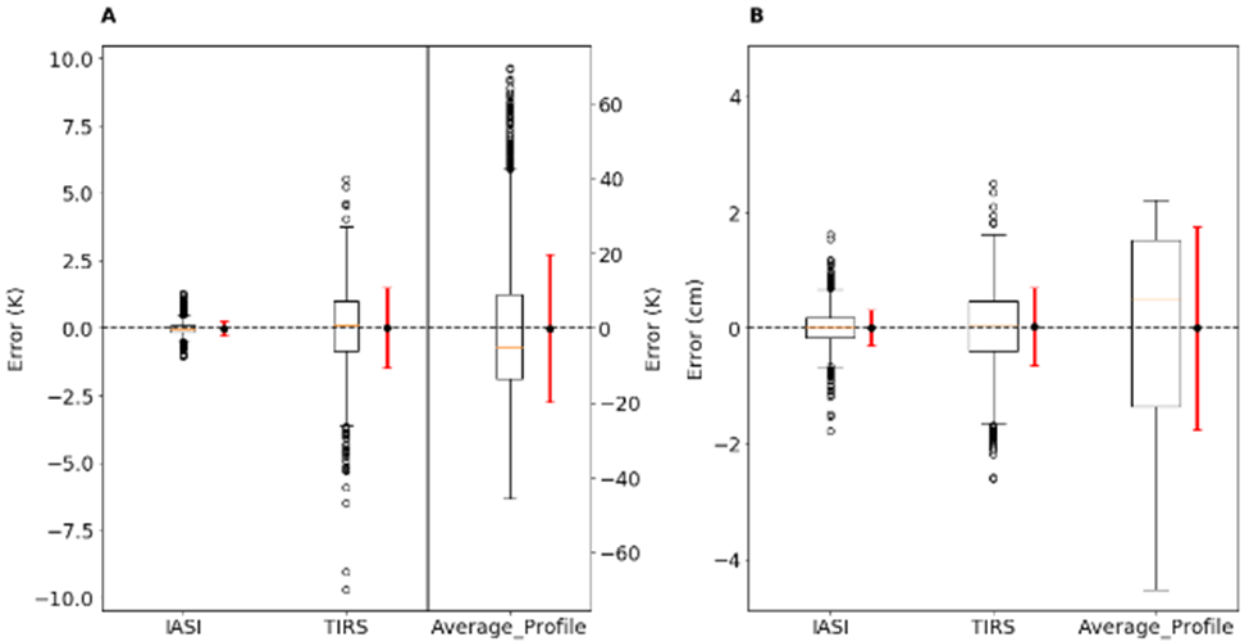


*Figure 3.2 – The error for skin temperature (A) and TPW (B) from the global base retrieval. The box plots show the spread of the data, while the red error bars indicate the error bias (represented by the black dot) +/– the standard deviation of error.*

Since the PREFIRE mission is primarily focused on the Arctic, it is also useful to characterize the retrieval performance on Arctic profiles separately. This was accomplished by analyzing the subset of the testing profiles located at latitudes north of 60°N. This subset contains 201 profiles. Figure 3.3 shows the retrieval bias and SDE in the Arctic for temperature and mixing ratio. Figure 3.4 shows the retrieval errors for skin temperature and total precipitable water in the Arctic. The magnitude of bias increased for all retrieved variables in the Arctic, meaning Arctic retrievals are less accurate. SDE also increased for Arctic retrievals of temperature and water vapor in all instruments. SDE showed a small decrease for the TIRS retrievals of skin temperature and total precipitable water, likely due to the smaller amount of atmospheric moisture typical of Arctic profiles. The increase in error is primarily seen at the surface and through the troposphere, where the most variability in atmospheric state occurs. However, this is true for all retrievals, not just for TIRS. The increase in bias and SDE is likely due to the distribution of profiles in the training data set. As previously mentioned in section

2.4.1, the SeeBor data set has more profiles in the midlatitudes and tropics than in the polar regions. As such, the base retrieval will be biased towards lower latitudes in terms of retrieval performance. In other words, since the algorithm had more training data in lower latitudes than at high latitudes, it is better equipped to perform retrievals at lower latitudes. That is not to say that the Arctic retrievals are inaccurate. As in the global case, the error for TIRS is between the two points of comparison, and is more similar to that of IASI than the average profile, meaning that TIRS is performing not only within the expected range of error, but with enough skill that it nearly comparable to a high spectral resolution instrument.

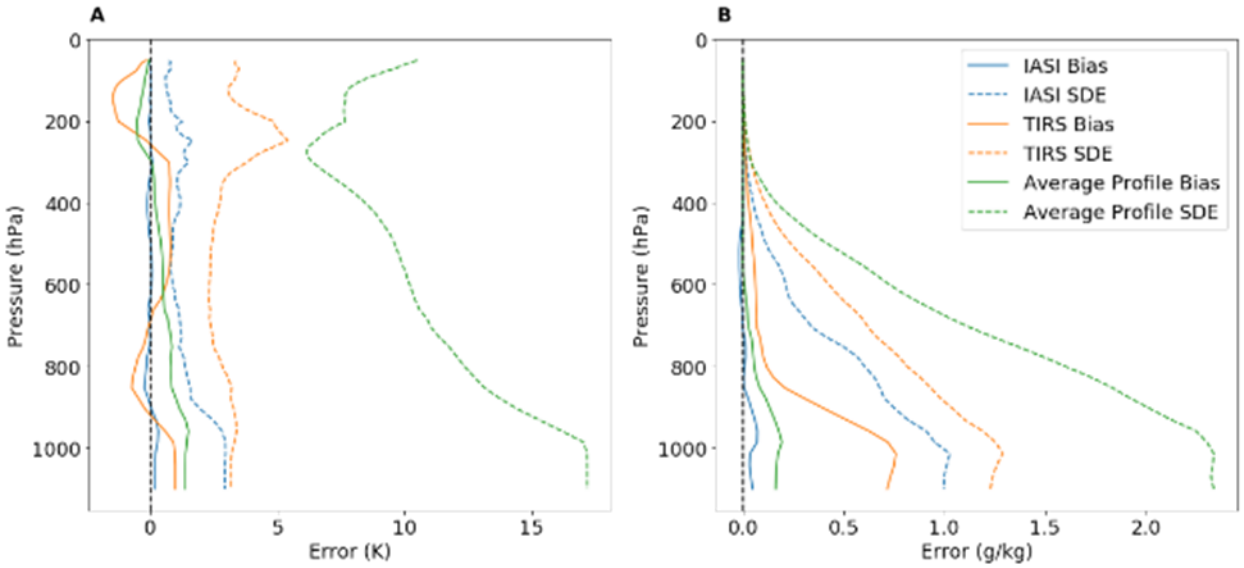


*Figure 3.3 – Arctic base retrieval performance for temperature (A) and mixing ratio (B). As with figure 3.1, solid color lines show retrieval bias, and dashed colored lines show SDE.*

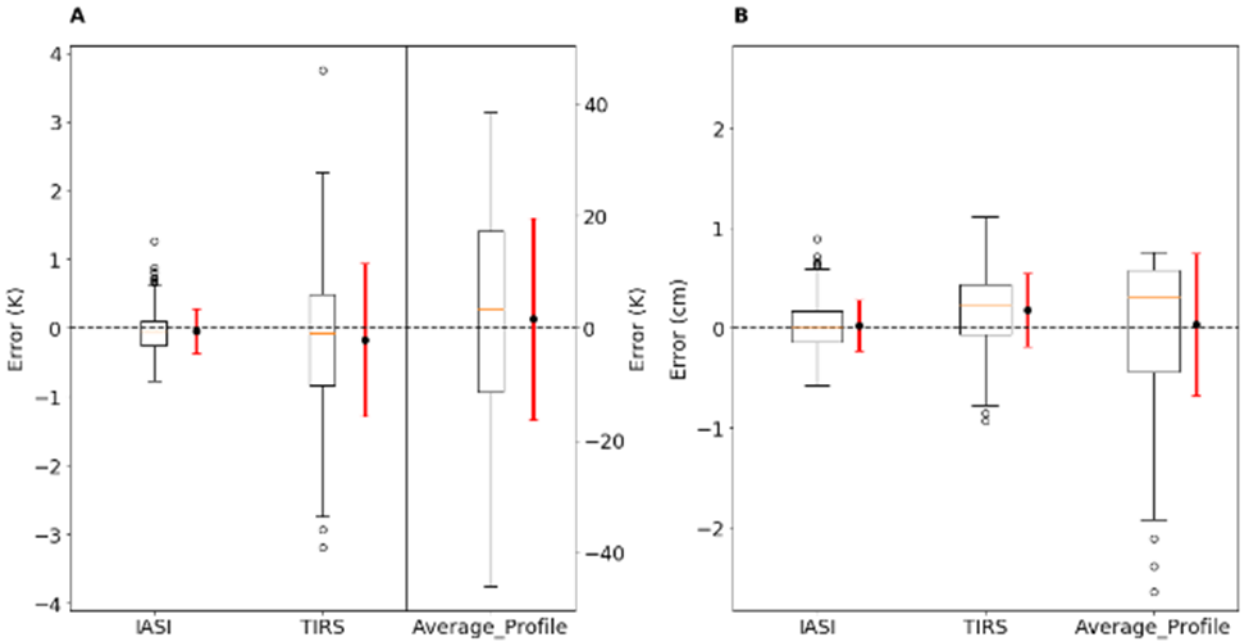


*Figure 3.4 – Arctic base retrieval performance for skin temperature (A) and total precipitable water (B). Box plots show median, interquartile range and spread of the retrieval error. To the right of the box plots, the error bars show the retrieval bias, represented the black dots, and the error bars show the bias +/– the SDE.*

## 3.2 Brightness Temperature Classification

One potential method for improving performance in the Arctic is to use a retrieval trained using profiles characteristic of colder environments. To accomplish this, brightness temperature classification (BTC) was implemented in the algorithm so it would be better tuned for specific environments and improve retrieval performance. To assess the change in retrieval performance, we look at the change in magnitude of the retrieval bias between the base and BTC retrievals. The change in retrieval bias for temperature and mixing ratio are shown in figure 3.5. In these graphs, and in the numbers reported, a negative difference in bias indicates a decrease in bias magnitude when BTC is implemented in the retrieval, bias magnitude being the absolute value of the retrieval bias. Averaged over the column, the bias magnitude difference was $1.43 * 10^{-2}$ K for temperature, and $-4.19 * 10^{-3}$ g/kg for mixing ratio. From these numbers, we see that when brightness temperature classification is implemented, bias increased for the temperature

measurements, and decreased for the measurements of moisture. However, these differences are small for all retrieved variables. As such we may conclude that the BTC retrieval displays similar accuracy to that of the unclassified retrieval.

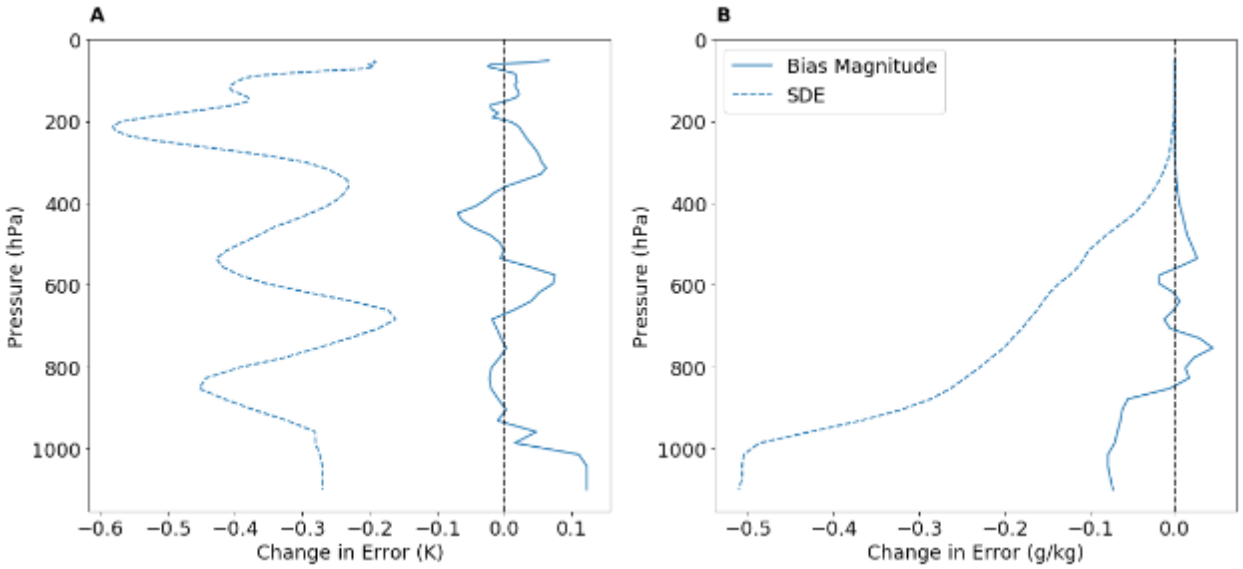


*Figure 3.5 – Change in error for the global retrieval between the base and BTC temperature retrieval (A) and mixing ratio retrieval (B). Solid blue line shows change in bias magnitude, dashed blue line shows change in SDE. Negative values indicate smaller errors in the BTC retrieval compared to the base retrieval.*

As the bias does not significantly change with implementation of BTC, change in retrieval performance will be realized through the change in SDE for the retrieved variables. The change in SDE between the base and BTC retrievals is evaluated as a percent difference in SDE:

$$\text{percent difference} = \frac{\text{SDE}_{\text{btc}} - \text{SDE}_{\text{base}}}{\text{SDE}_{\text{base}}} * 100$$

This quantity is useful for assessing changes in retrieval error because it is independent of the magnitude of the original variable. The reason for not assessing percent difference when analyzing change in bias is because bias is ideally centered around zero, so even a small change in bias may result in a large percent difference that does not accurately represent the results. This is not the case for SDE, which is not equal to zero, so the percent difference is a reasonable means of demonstrating the change in error.

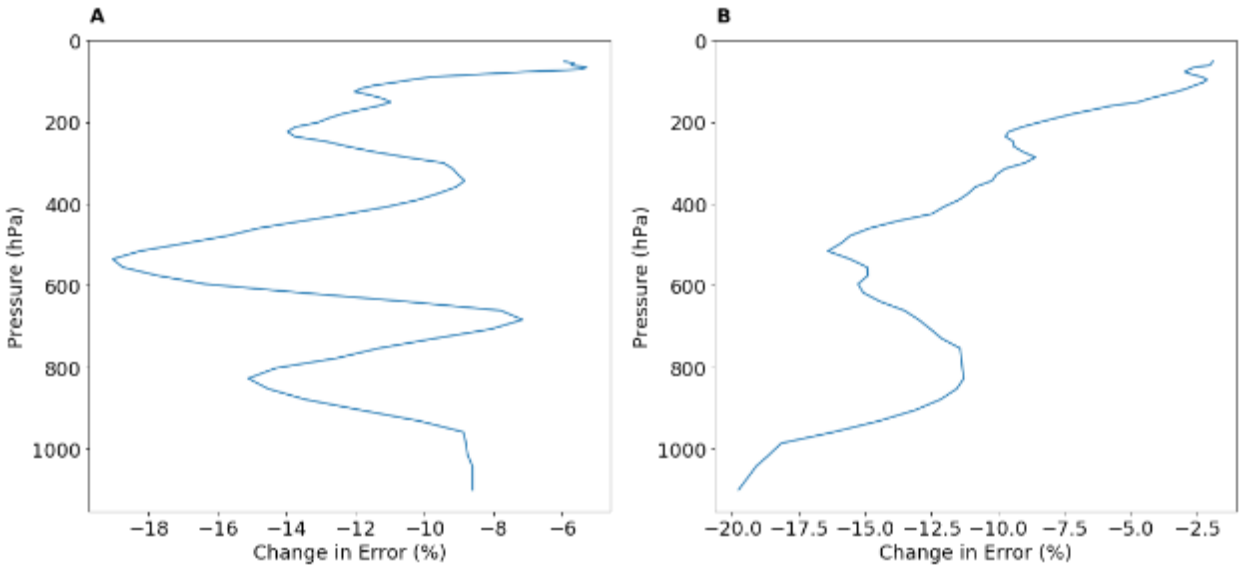


*Figure 3.6 – Percent change in SDE for the global temperature (A) and mixing ratio (B) retrievals when BTC is applied. Negative values indicate a lower error in the BTC retrieval compared to the base retrieval.*

The percent difference of SDE in the global case is shown in Figure 3.6. As with the change in bias, a negative difference indicates a decrease in error when BTC is applied. Therefore, a negative percent difference in SDE represents an increase in retrieval precision. The temperature SDE decreased overall, with the largest improvement in retrieval accuracy in the middle of the troposphere, between 500 and 600 hPa, with a peak difference of –19%. Mixing ratio likewise showed decreases in SDE throughout the column when BTC is applied. Figure 3.6b shows distinct peaks at 500 hPa and at the surface, around 1000 hPa. The largest decrease in mixing ratio SDE was located near 1000 hPa, where SDE decreased by 20%. The average change in SDE for temperature is –10% over the column. For mixing ratio, the average percent difference is –7.8% over the column. As for the non-profile variables, skin temperature SDE decreased by 40%, and TPW SDE had a percent difference of –21%. Thus, even though bias did not change, we see that the effect of applying brightness temperature classification to the retrieval is reduction in the SDE, and improvement in retrieval precision.

The Arctic saw more improvement with the addition of BTC than the total global test set. The magnitude of temperature bias (Figure 3.7a) generally decreased over our atmospheric region of interest, with an average change in magnitude of $-1.75*10^{-2}$ K. The large changes above 10 hPa are not significant, again due to the lack of signal at very high altitudes. As for mixing ratio (Figure 3.7b), there was also an overall decrease in bias magnitude, with the largest changes occurring near the surface, and an average change in bias of $-3.27*10^{-2}$ g/kg. The SDE for these variables likewise saw more improvement. The percent difference of temperature and mixing ratio SDE in the Arctic is shown in figure 3.8. The temperature SDE had the largest percent difference around 800 hPa and 200 hPa. The maximum difference in SDE was –25% at just below the 800 hPa pressure level. The average percent difference in SDE over the column was –8.6%. The mixing ratio SDE behaved similarly in the arctic as it did in the global case: the SDE decreased over the entire column, with the largest difference near 1000 hPa, with a maximum reduction in SDE of 37 %. The average change in mixing ratio SDE over the column was –15%. As for skin temperature and TPW, skin temperature bias in the Arctic increased by $9.83*10^{-2}$K, though the SDE decreased by 25%. TPW bias decreased by –0.11cm, and the SDE decreased by 43%.

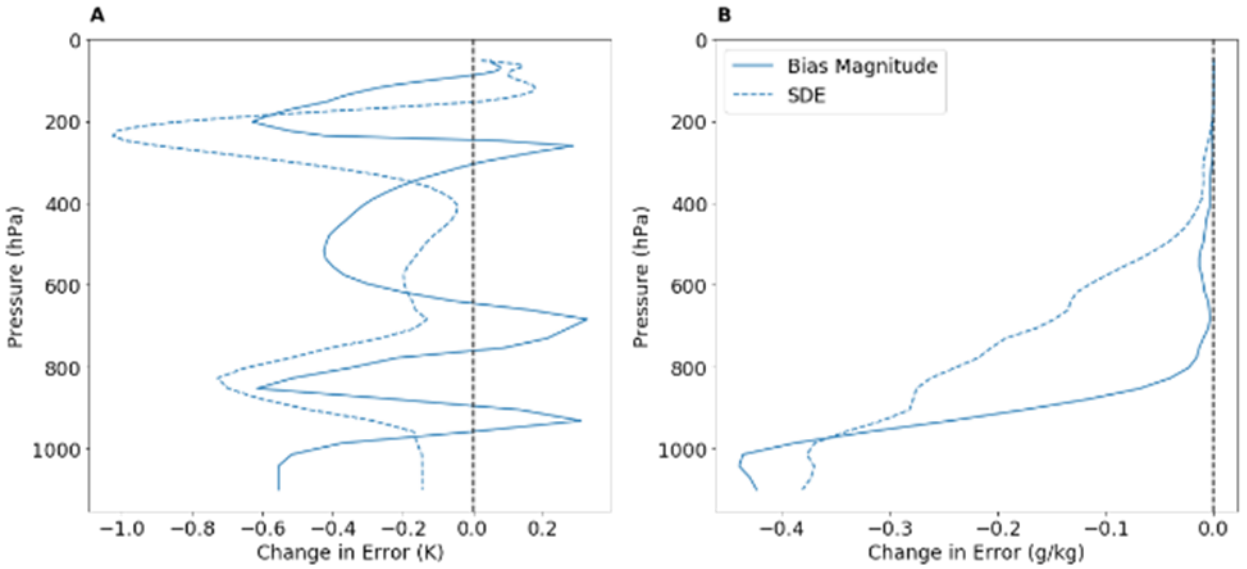


*Figure 3.7 – Change in error for the Arctic retrieval between the base and BTC temperature retrieval (A) and mixing ratio retrieval (B). Solid blue line shows change in bias magnitude, dashed blue line shows change in SDE. Negative values indicate smaller errors in the BTC retrieval compared to the base retrieval.*

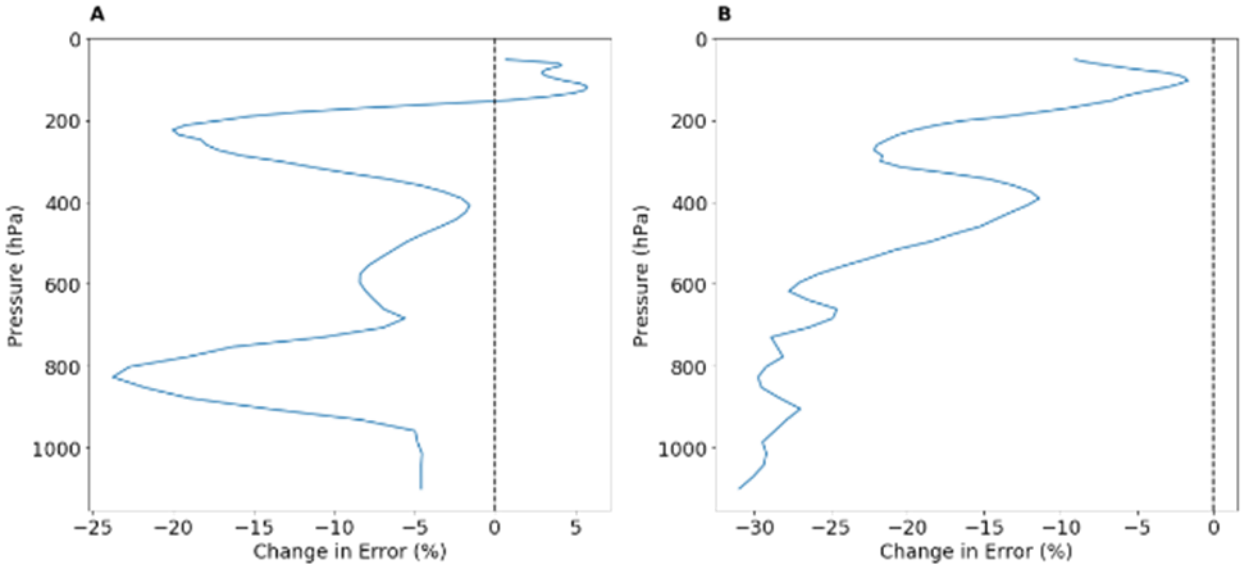


*Figure 3.8 – Percent change in SDE for the Arctic temperature (A) and mixing ratio (B) retrievals when BTC is applied. Negative values indicate a lower error in the BTC retrieval compared to the base retrieval.*

The reasoning behind the improvements in Arctic retrievals with the addition of BTC harkens back to the spatial distribution of the test data. Since there are more profiles in warmer

regions, Arctic retrievals had larger errors than the overall global case. By implementing BTC, a retrieval in the Arctic is performed with a different set of regression coefficients than a retrieval in the tropics. Therefore, a cold dry profile will be retrieved using regression coefficients that more closely represent it, hence the large reduction of errors in the Arctic retrievals.

## 3.3 Retrieval with and without FIR

As the PREFIRE mission will be illustrating the benefit of regular measurements of the FIR from space, it would be remiss to not highlight how the FIR measurements of TIRS benefit the atmospheric retrievals. In order to do this, it is necessary to compare a retrieval with the FIR included to a retrieval without FIR measurements. This is done in a similar respect in section 3.1, where the TIRS retrieval is compared to that of IASI, however, it is not a direct comparison due to differences in channels and spectral resolution between the two instruments. Here, a more direct comparison is performed by conducting the TIRS retrieval with only the mid infrared channels and comparing it to the full TIRS retrieval.

Only 11 of the 54 TIRS channels are found in the MIR. Therefore, the regression coefficients for the MIR subset were calculated using 12 principal components, the additional PC coming from surface pressure. Regression coefficients for the full spectra were also recalculated with 12 PCs so that the comparison only shows the difference between the amount of information, rather than the differing number of channels. Brightness temperature classification was used in these retrievals to mimic the anticipated implementation of the algorithm with real data.

As done in the analysis of brightness temperature classification, the comparison of the two retrievals is performed by assessment of the difference in retrieval bias magnitude as well as the percent difference in SDE.

In the global case, the change in bias between the full retrieval and the MIR retrieval is negligible. The difference in bias magnitude for temperature and mixing ratio are shown in figure 3.9. The changes in temperature bias are small, with an average difference of $-8.7*10^{-4}$ K over the column. Such small differences are likely due to random noise rather than an actual change in retrieval bias. The mixing ratio similarly had changes in bias that were likely due to differences in noise. Below 853 hPa, the change in bias is negative, indicating a reduction in bias magnitude when FIR measurements are included. Above 853 hPa, the bias is primarily positive. However, the average change in bias over the column is small: $9.5*10^{-4}$ g/kg. For the non-profile variables, the change in bias was likewise small. Skin temperature bias changed by $-9.2*10^{-3}$K, and total precipitable water bias by $-1.4*10^{-3}$cm when FIR measurements are included.

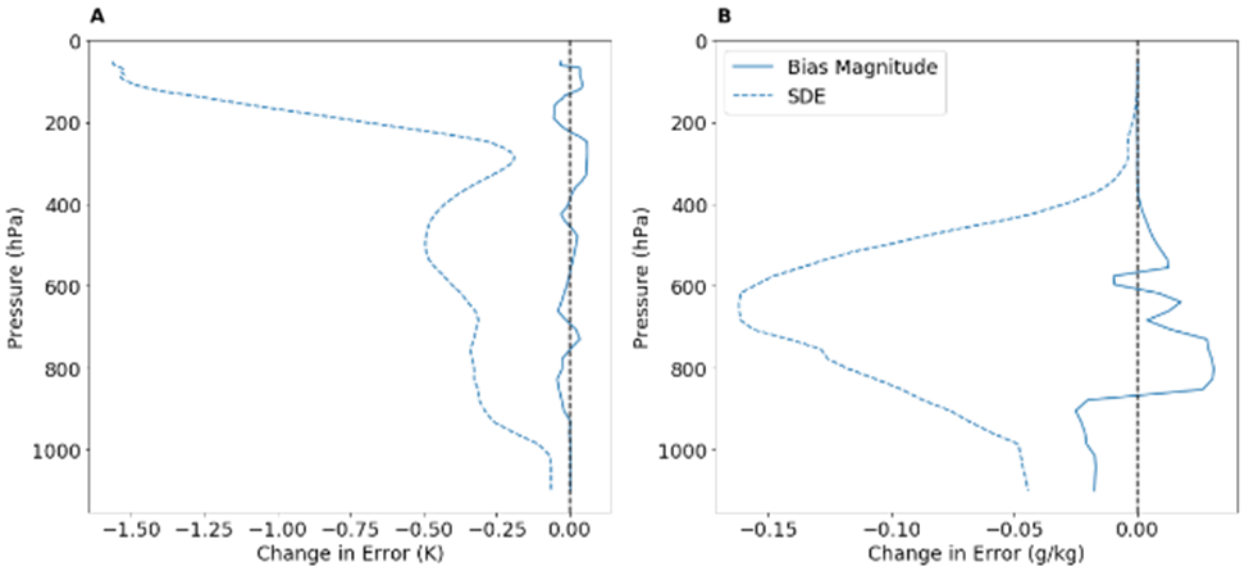


*Figure 3.9 – Difference in error between the MIR only retrieval and all-channel retrieval for temperature (A) and mixing ratio (B), using all profiles in the validation dataset. The solid blue line shows change in bias magnitude. The dashed blue line shows change in SDE. Negative values indicate the all-channel retrieval had lower error than the MIR-only retrieval.*

It is in the SDE where we begin to see the benefit of FIR measurements in retrievals. The percent difference in global temperature SDE is shown in figure 3.10a, where the negative values indicate an overall reduction in SDE when FIR measurements are included. The average

difference in temperature SDE is –16% with a peak reduction in error of –34% at 47 hPa. Two other peaks of note include one at 118 hPa with a reduction of 32% and a smaller peak in the troposphere at 516 hPa, where the reduction of SDE was –21%. The mixing ratio retrieval also saw a reduction of error in the global case, as demonstrated in Figure 3.10b. The maximum reduction of error is found at 516 hPa, with percent difference in SDE of –18%. A secondary peak exists at 212 hPa, where the reduction in error reaches –11%. The average change in mixing ratio SDE in the global case is –5.8%. Skin temperature and TPW also decreased, with a skin temperature SDE changing by –2.3%, and TPW SDE by –9.5%. These results suggests that the addition of FIR measurements influence the retrieval precision by reduction of the SDE in the global case.

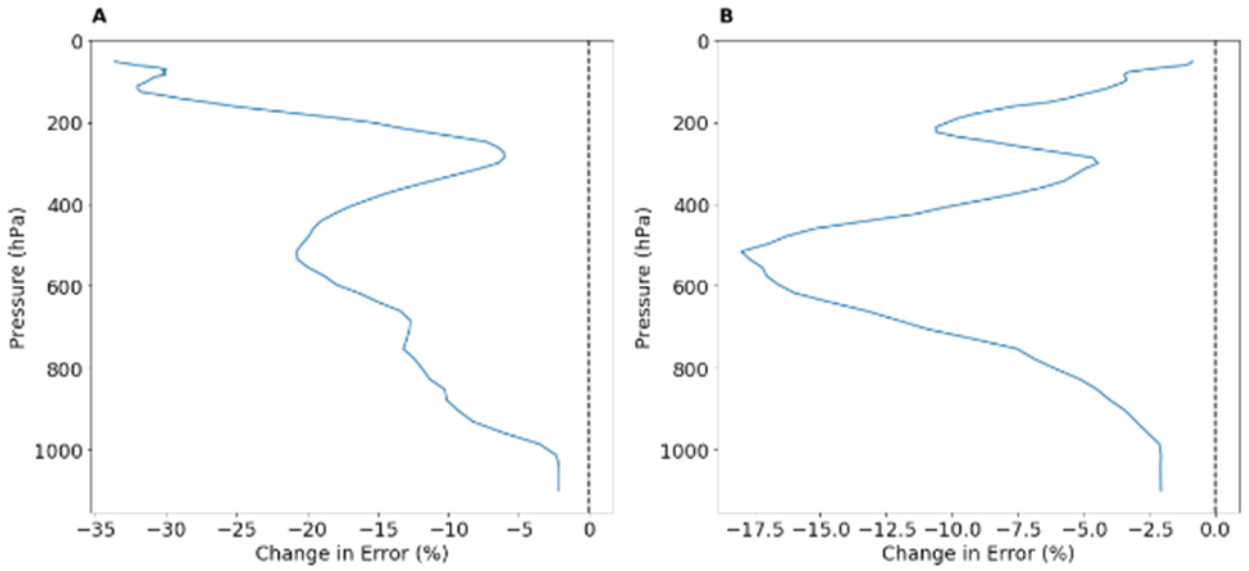


*Figure 3.10 – Percent difference in SDE between the MIR-only and all-channel retrievals for temperature (A) and mixing ratio (B) using all profiles in the testing dataset. Negative values indicate a smaller error in the all-channel retrieval compared to the MIR only retrieval.*

The Arctic case is a different story, however. The change in temperature bias, shown in figure 3.11a, is primarily negative, meaning the retrieval improved in the Arctic when FIR channels were included. This is likely due to the drier atmosphere where the FIR can see down to

the surface. The largest reduction in bias magnitude were located at 497 hPa where bias decreased by –0.4976 K. The maximum reduction occurs in the stratosphere around 83 hPa, where the change in bias was –1.919K. There are a few places where bias increased when the FIR measurements are included: 932 hPa and 260 hPa, where bias increased by $8.492*10^{-2}$K and 0.2309K respectively. However, overall, there was an average reduction in bias of –0.8824K over the column. Likewise, the mixing ratio bias saw a net reduction in the Arctic with the addition of FIR measurements, as shown in figure 3.11b. There is a small region in the troposphere in which bias increased, with a peak at 779 hPa, where bias increased by $1.9*10^{-2}$ g/kg. Bias decreased through the rest of the column, primarily near the surface with a maximum decrease in bias of $-8.7*10^{-2}$ g/kg. The average change in bias was $-5.7*10^{-3}$ g/kg. As for the non-profile variables, the skin temperature bias changed by –0.88K and TPW by $-9.3*10^{-3}$cm.

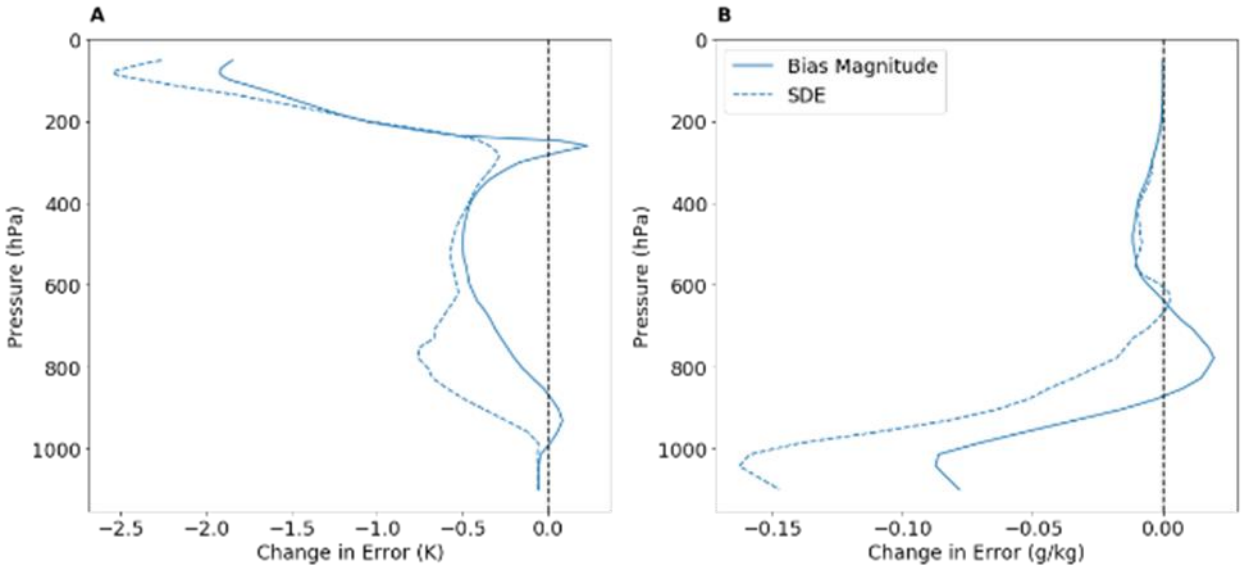


*Figure 3.11 – Difference in error between the MIR only retrieval and all-channel retrieval for temperature (A) and mixing ratio (B), using only Arctic profiles from the validation dataset. The solid blue line shows change in bias magnitude. The dashed blue line shows change in SDE. Negative values indicate the all-channel retrieval had lower error than the MIR-only retrieval.*

As with the global case, the SDE was reduced more than bias in the Arctic case when FIR measurements were included. The change in SDE for temperature is shown in figure 3.12. The change in SDE in the Arctic has a similar shape to the change in bias curve. On average, the temperature SDE decreased by 20%, with a peak reduction of –46% in the stratosphere at around 80 hPa. The mixing ratio SDE behaves similarly to bias, in that there is a region in the troposphere (here at 639 hPa), where the SDE increased, though by less than 1%, but the change in SDE is negative through the rest of the column. The largest decrease in error occurs at the lowest levels with a peak error reduction of around 15%. There is a similar peak in the upper troposphere at 286 hPa, where the reduction in SDE is –13%. The average change in mixing ratio SDE is –6.2%. For both temperature and mixing ratio, the change in error is greater in the Arctic case than in the global. This is also true for skin temperature and TPW. Skin temperature SDE changed by –12% and TPW SDE by –10% in the Arctic case.

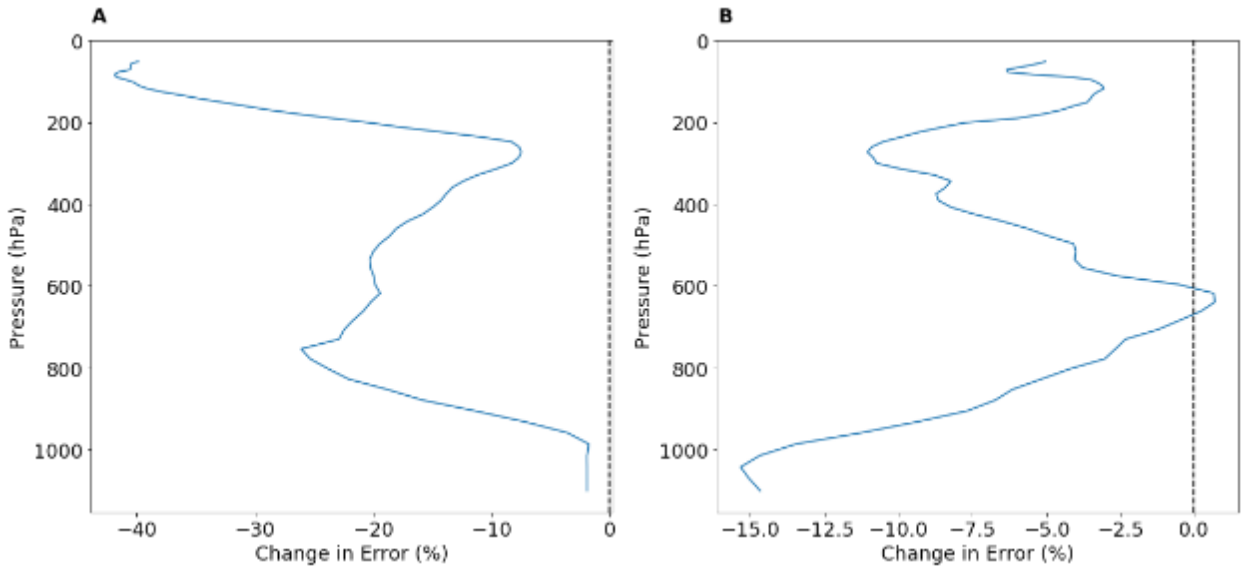


*Figure 3.12 – Percent difference in SDE between the MIR-only and all-channel retrievals for temperature (A) and mixing ratio (B) using only Arctic profiles from the testing dataset. Negative values indicate a smaller error in the all-channel retrieval compared to the MIR only retrieval.*

# Chapter 4: Conclusions

This paper introduces a PCR temperature and humidity retrieval for the TIRS instrument aboard PREFIRE. Three different trials of the PCR retrieval algorithm were conducted to illustrate its performance relative a high spectral resolution instrument and assess the impact of preconditioning the retrieval and the influence of FIR wavelengths. The average testing profile serves as a baseline or lower limit to the skill of the retrieval, and the IASI retrieval acts as an upper limit, being a retrieval from an instrument with a higher spectral resolution than TIRS.

In the initial test, the TIRS retrieval performance was between the other two retrievals, with a bias generally consistent with zero, and a SDE between that of IASI and the average training profile. The average standard deviation of temperature error in troposphere was 1.3 K for IASI, 2.8K for TIRS, and 12.8 K for the average training profile. For water vapor, the tropospheric average of SDE was 0.65 g/kg for IASI, 1.1 for TIRS, and 2.2 for the average profile. The primary purpose of this first test was to establish a baseline for the TIRS retrieval performance. From the results we conclude that the TIRS retrieval performance is within expected limits, and performs more similarly to the high spectral resolution instrument represented by IASI, rather than the no-skill benchmark of the average training profile.

Brightness temperature classification was found to have a positive impact on the PREFIRE PCR retrieval. Brightness temperature classification uses the brightness temperature of a certain channel (in this case, the 10.97 μm channel) as a proxy for surface temperature, and selects a set of retrieval coefficients based on that brightness temperature. This method has the potential to reduce errors in the retrieval from skewed representation of surface temperatures in the training data set. In the case of the SeeBor dataset used to train the PREFIRE retrieval algorithm, more profiles originate from mid and lower latitudes than high latitudes. Therefore, use of brightness temperature classification reduced the error in high-latitude profiles, which had

lower representation. Findings showed that BTC resulted in a reduction in SDE throughout the atmosphere for the temperature retrieval, and reduction in SDE at lower altitudes for the mixing ratio retrieval. The changes in retrieval bias were small and may be attributed to random noise rather than improved retrieval bias. More improvement from BTC could be seen in the Arctic compared to the overall global case. Changes in both bias and SDE occurred in the Arctic retrievals, primarily showing improvement in retrieval performance. Though in some regions the bias magnitude increased with the added BTC, these positive changes are small and may be attributed to random noise. This affirms the assertion that the BTC would help to compensate for the lower representation of high-latitude profiles in the training dataset.

The third test was a comparison of the TIRS retrieval with and without far infrared channels. By reducing the number of principal components to equal the number of mid infrared channels, this comparison is made possible. This comparison allows us to see how measurement of FIR impacts the ability to retrieve atmospheric variables such as temperature and water vapor. Overall, adding the FIR channels to the retrieval saw minimal reduction in bias from the MIR-only retrieval, but significant reduction in the standard deviation of error. More improvement was seen in the temperature retrieval than in the mixing ratio retrieval. Larger improvements were seen when considering only profiles from the Arctic. In both the temperature and mixing ratio retrievals, as well as the skin temperature and TPW retrievals, the addition of FIR channels resulted in decreased retrieval bias and SDE. As with the comparison to IASI in the base retrieval test, this third test confirms the benefit of using FIR measurements in atmospheric retrievals, particularly in cold and dry regions such as the Arctic.

# Chapter 5: Future Work

Though the PREFIRE PCR retrieval algorithm is fully functional, there are still ways in which the algorithm may be built upon and improved. Prior to the launch of the PREFIRE satellites, the current PCR algorithm will be retrained with supplementary profiles to further improve retrieval performance. In particular, additional profiles from high latitude regions will be included in the training data to account for skewed representation of atmospheric states in the SeeBor dataset.

In addition, the algorithm will be tested with cloudy scenes to assess the retrieval performance in the presence of clouds. Given the sensitivity of the FIR to water and water vapor, the PCR algorithm could potentially serve as a test for clouds in a scene. Moreover, even if the retrieval is not accurate within and below a cloud base, information of the atmosphere above the cloud may still be obtained.

# Bibliography

Ackerman, T. P., & Stokes, G. M., (2003). The Atmospheric Radiation Measurement Program. *Physics Today*, *56*, 38-44. doi:10.1063/1.1554135.

Borbas, E. E., Knuteson, R.O., Seemann, S. W., Weisz, E., Moy, L., & Huang, H-L. (2007). A high spectral resolution global land surface emissivity database. Poster presented at the Joint 2007 EUMETSAT Meteorological Satellite 15th AMS Satellite Meteorology and Oceanography Conference, Amsterdam, The Netherlands, 24-28 September 2007.

Borbas, E., Seemann, S. W., Huang. H-L., Li, J., & Menzel, W. P. (2005). Global profile training database for satellite regression retrievals with estimates of skin temperature and emissivity. Proc. Int. ATOVS Study Conf. XIV, Beijing, China, CIMSS/University of Wisconsin—Madison, 763–770.

Canas, T., Murray, J., & Harries, J. (1997). Tropospheric airborne Fourier transform spectrometer (TAFTS). *Proceedings of SPIE 3220, Satellite Remote Sensing of Coulds and the Atmosphere*, doi: 10.1117/12.301139

Harries, J., Carli, B., Rizzi, R., Serio, C., Mlynczak, L., Palchetti, T., & et al. (2008). The Far-Infrared Earth. *Rev. Geophys, 46*.

Hilton, F., and coauthors (2012). Hyperspectral Earth Observation from IASI: Five Years of Accomplishments. *Bulletin of the American Meteorological Society*, *93(3)*, 347-370.

Huang, H. L., Chen, X. H., Zhou, D. K., & Liu, X. (2016). An observationally based global band-by-band surface emissivity dataset for climate and weather simulations. *Journal of the Atmospheric Sciences, 73,* 3541 – 3555. doi:10.1175/JAS-D-15-0355.1

L'Ecuyer, T. S., Drouin, B. J., Anheuser, J., Grames, M., Henderson, D. S., Huang, X., et al. (2021). The Polar radiant energy in the far infrared experiment. *Bulletin of the American Meteorological Society, 102*(7), 1431-1449.

Libois, Q., Ivanescu, L., Blanchet, J.-P., Schulz, H., Bozem, H., Leaitch, W. R., et al. (2016). Airborne observations of far-infrared upwelling radiance in the Arctic. *Atmospheric Chemistry and Physics, 16*, 15689-15707. doi:10.5194/acp-16-15689-2016.

Liu, X., Smith, W. L., Zhou, D. K., & Larar, A. (2006). Principal Component – Based Radiative Transfer Model for Hyperspectral Sensors: Theoretical Concept. *Applied Optics 45(1),* 201-209. doi: 10.1364/AO.45.000201

Merrelli, A., & Turner, D. (2012). Comparing information content of upwelling far-infrared and midinfrared radiance spectra for clear atmosphere profiling. *Journal*

*of Atmospheric and Oceanic Technology*, *29*, 510-526. doi: 10.1175/JTECH-D-11-00113.1.

Mlawer, E. J., Turner, D. D., Paine, S. N., Palchetti, L., Bianchini, G., Payne, V. H., et al. (2019). Analysis of Water Vapor Absorption in the Far-Infrared and Submillimeter Regions Using Surface Radiometric Measurements From Extremely Dry Locations. Journal of Geophysical Research: Atmospheres, 124(14), 8134–8160. https://doi.org/10.1029/2018JD029508

Mlynczak, M., Harries, J., Rizzi, R., Stackhouse, P., Kratz, D., Johnson, D., & et al. (2002). Far-infrared: a frontier in remote sensing of Earth's climate and energy balance. *International Symposium on Optical Science and Technology*, doi: 10.1117/12.454247

Nalli, N., Minnet, P., & Van Delst, P. (2008). Emissivity and reflection model for calculating unpolarized isotropic water surface-leaving radiance in the infrared. I: Theoretical development and calculations. *Applied Optics 47,* 3701-3721.

Palchetti, L., Bianchini, G., Pellegrini, M., Esposito, F., Restieri, R., Pavese, G. (2004). Radiometric performances of the Fourier transform spectrometer for the Radiation Explorer in the Far-Infrared (REFIR) space mission. *Proceedings of SPIE, Sensors, Systems, and Next-Generation Satellites VIII.* doi: 10.1117/12.565548

Palchetti, L., Brindley, H., Bantges, R., Buehler, S., Camy-Peyret, C., Carli, B., et al. (2020). FORUM: Unique far-infrared satellite observations to better understand how Earth radiates energy to space. *Bulletin of the American Meteorlogical Society, 101*(12), E2030-E2046. doi: 10.1175/BAMS-D-19-0322.1.

Rizzi, R., Serio, C., & Armorati, R. (2002). Sensitivity of broadband and spectral measurements of outgoing radiance to changes in water vapor content. *Proceedings of SPIE, International Symposium on Optical Science and Technology*, doi: 10.1117/12.454250).

Seemann, S.W., Borbas, E. E., Knuteson, R. O., Stephenson, R. O., Huang, H.-L. (2007). Development of a Global Infrared Land Surface Emissivity Database for Application to Clear Sky Sounding Retrievals from Multi-spectral Satellite Radiance Measurements. *Journal of Applied Meteorology and Climatology*, *47*, 108-123.

Shahabadi, M. B., & Huang, Y. (2014). Measuring Stratospheric H2O with an airborne spectrometer. *Journal of Atmospheric and Oceanic Technology, 31*, 1502-1515. doi: 10.1175/JTECH-D-13-00191.1.

Shine, K., Ptashnik, I., & Rädel, G. (2012). The water vapor continuum: Brief history and recent developments. *Survey of Geophysics, 33*, 535-555. doi: 10.1007/s10712-011-9170-y.

Spankuch, D. & Döhler, W. (1985). Radiative properties of cirrus clouds in the middle IR derived from Fourier spectrometer measurements from space. *Zeitschrift für Meteorologie, 35 (6)*. 314-324

Timofeev, Y.M., Polyakov, A.V., Kozlov, D.A., Zavelevich, F.S., Golovin, Y.M., Döhler, W., Oertel, D., & Spankuch, D. (2019). Comparison between the spectra of outgoing infrared radiation for different years. *Izvestiya, Atmospheric and Oceanic Physics, 55,* 956-962. doi: 10.1134/S0001433819090524

Turner, D.D., & Mlawer, E. J. (2010). Radiative Heating in Underexplored Bands Campaigns (RHUBC). *Bulletin of the American Meteorological Society, 91(7),* 911-924. doi: 10.1175/2010BAMS2904.1

Warwick, L., Brindley, H., Di Roma, A., Fox, S., Havemann, S., Murray, J., et al. (2022). Retrieval of Tropospheric Water Vapor from Airborne Far-Infrared Measurements: A Case Study. *Journal of Geophysical Research: Atmospheres, 127 (7),* doi: 10.1029/2020JD034229

Weisz, E., Huang, H.-L., Li, J., Borbas, E., Baggett, K., Thapliayal, P.K., Guan, L. (2007). International MODIS and AIRS processing package: AIRS products and applications. *Journal of Applied Remote Sensing, 1*.

Wielicki, B., Young, D.F., Mlynczak, M. G., Thome, K.J., Leroy, S., Corliss, J., et al., (2013). Achieving Climate Change Absolute Accuracy in Orbit. *Bullet of the*

*American Meteorological Society, 94(10).* 1591-1539. doi: 10.1175/BAMS-D-12-00149.1