

REPORT

Earth Observation: Methods & Applications

Atmospheric sounding using satellite microwave observations

Lecturer: Dr. Filipe Aires

Student: Mai Nhu Tin (SA, USTH)



Department of Space and Applications
University of Science and Technology of Hanoi

Contents

1	Introduction	1
1.1	Generalisties	1
1.2	AMSU-A and –B measurements	1
2	Dataset	2
2.1	Geophysical temperature	2
2.2	Geophysical humidity	3
2.3	AMSU	4
3	Correlation	5
4	Modelization	6
5	Classification	9
6	Principle Components Analysis	12

1 Introduction

1.1 Generalities

The Advanced Microwave Sounding Units (AMSU) A and B, on NOAA satellites, play a crucial role in measuring radiation from the atmosphere and Earth's surface. AMSU-A focuses on gathering atmospheric temperature data from about 3 hPa (45km) down to the surface, while AMSU-B targets water vapor in the atmosphere at 183 GHz. This dual capability makes AMSU data valuable for Numerical Weather Prediction (NWP) models, providing accurate monitoring of air temperature and moisture with good temporal and spatial coverage. Unlike infrared observations, AMSU data are less affected by certain types of clouds.

Several methods have been developed to extract temperature and humidity profiles from AMSU data, including surface and atmosphere modeling, passive microwave measurements, and neural network techniques.

While AMSU data over the ocean are routinely used in NWP systems, their full potential over land remains untapped due to challenges in distinguishing between surface and atmospheric signals. Efforts have been made to integrate AMSU data over land, but accurately estimating land emissivity poses a significant obstacle.

This study aims to evaluate the feasibility of using all AMSU-A and -B channels over the ocean to obtain temperature and moisture profiles down to the surface. The goal is to improve our understanding and utilization of satellite observations for weather forecasting and climate monitoring.

1.2 AMSU-A and -B measurements

The AMSU instrument has been active on NOAA satellites since 1998. AMSU-A has 12 channels focused near oxygen absorption lines below 60 GHz and four window channels at 23.8, 31.4, 50.3, and 89 GHz. AMSU-B includes two window channels at 89 and 150 GHz, and three channels targeting the 183.31 GHz water vapor line. These instruments have different viewing angles and sample Earth views differently. They observe a wide range of angles, from -48° to $+48^\circ$, with a maximum local zenith angle of 58° . The characteristics of each channel are listed in Figure 2. This study focuses on analyzing AMSU data over the ocean under clear sky conditions.

Figure 1 displays weighting functions for all AMSU channels, calculated for a standard tropical atmosphere using a radiative transfer model. These functions show how much each layer of the atmosphere contributes to the measured radiation. The peak altitude in the weighting function increases with higher viewing angles due to longer paths through the atmosphere. Window channels show peaks closer to the surface, indicating that most of the radiation measured by these channels comes from the surface and the lower atmosphere. These channels are useful for estimating parameters like water vapor, precipitation, or cloud liquid water over the ocean. Accurately estimating surface emissions is crucial for separating their effects from atmospheric signals.

AMSU observes the Earth from various angles, which makes developing retrieval algorithms tricky. Some methods adjust observations to simulate measurements taken directly below the satellite. However, in our study, we assume all observations are made directly below the satellite without any adjustments.

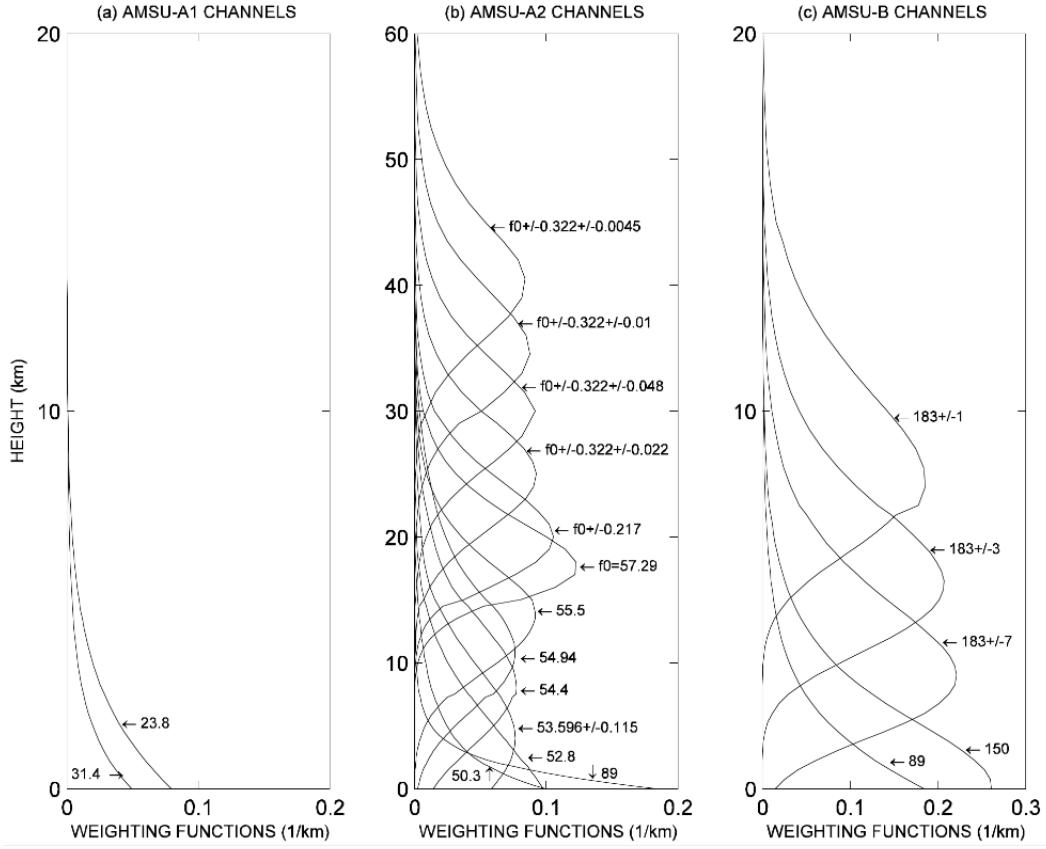


Figure 1: AMSU-A and -B weighting functions for a US standard tropical atmosphere ($\text{WV} = 42 \text{ kg/m}^2$) at nadir, assuming a surface temperature of 299 K and a surface emissivity of 0.95 for: (a) AMSU-A1 (b) AMSU-A2, and (c) AMSU-B channels.

2 Dataset

The dataset contains three type of measurements on ocean temperature and humidity all over the world. These include the geophysical data, which describes the in situ measurements on Earth at 43 vertical layers are numbered from the Top of Atmosphere (TOA) down to the surface (from 1-4: mesosphere, from 5-19: stratosphere, rest: troposphere), we refer this as the true physical parameters. And space data from two instruments embedded in weather satellites, AMSU-A and AMSU-B in 15 and 5 channels as describes in Figure 1, respectively.

2.1 Geophysical temperature

Figure 3 shows the maps of temperature from latitude -77 to 90 and longitude -180 to 180 in six layers at different heights. We can see clearly that the temperature distribution varies as we move from the TOA down to the surface. At the surface, the temperature is highest across the equator and decreasing as going to the two poles. The opposite pattern occur at the TOA when the temperature decreases gradually from the North Pole to the South Pole indicates that this might be summer in the northern hemisphere at the time taking the measurements.

More details about vertical temperature profile in all layers are shown in Figure 4. In the left plot, seven samples are chose with the same longitude at the middle of Atlantic Ocean in order to remove the effect of missing continent data. We can see the vertical

Channel No	Frequency (GHz)	Noise equivalent (K)	Resolution at nadir (km)
AMSU-A			
1	23.8	0.20	48
2	31.4	0.27	48
3	50.3	0.22	48
4	52.8	0.15	48
5	53.596+/-0.115	0.15	48
6	54.4	0.13	48
7	54.9	0.14	48
8	55.5	0.14	48
9	57.290=f ₀	0.20	48
10	f ₀ +/- 0.217	0.22	48
11	f ₀ +/- 0.322 +/- 0.048	0.24	48
12	f ₀ +/- 0.322 +/- 0.022	0.35	48
13	f ₀ +/- 0.322 +/- 0.010	0.47	48
14	f ₀ +/- 0.322 +/- 0.0045	0.78	48
15	89	0.11	48
AMSU-B			
16	89	0.37	16
17	150	0.84	16
18	183.31 +/- 1	1.06	16
19	183.31 +/- 3	0.70	16
20	183.31 +/- 7	0.60	16

Figure 2: AMSU-A and -B channel description.

shape of the atmosphere up to mesosphere. The temperature decreases with height in the troposphere, decreasing slowly or even constant in the stratosphere then increasing until it meets the stratopause. One also recognize the temperature at the surface are nearly the same for regions from the Arctic Circles to the Equator but starting increase with height from the stratosphere, highest at the North Pole and lowest at the South Pole as more illuminating by the Sun in the northern hemisphere. The right plot shows the profile across the equator, the shape is similar in all sample longitudes, only a slightly difference in temperature in the mesosphere. It is also very interesting if we look into the shape of the profile at the regions near the equator, there are no stable sublayer where the temperature is constant with height but instead the temperature suddenly inverses its trend at the pressure of 100 hPa.

2.2 Geophysical humidity

The water vapour almost absents in mesosphere and stratosphere, which leads to very low humidity percentage in these layers as seen in Figure 5. Humidity appears when reaching the

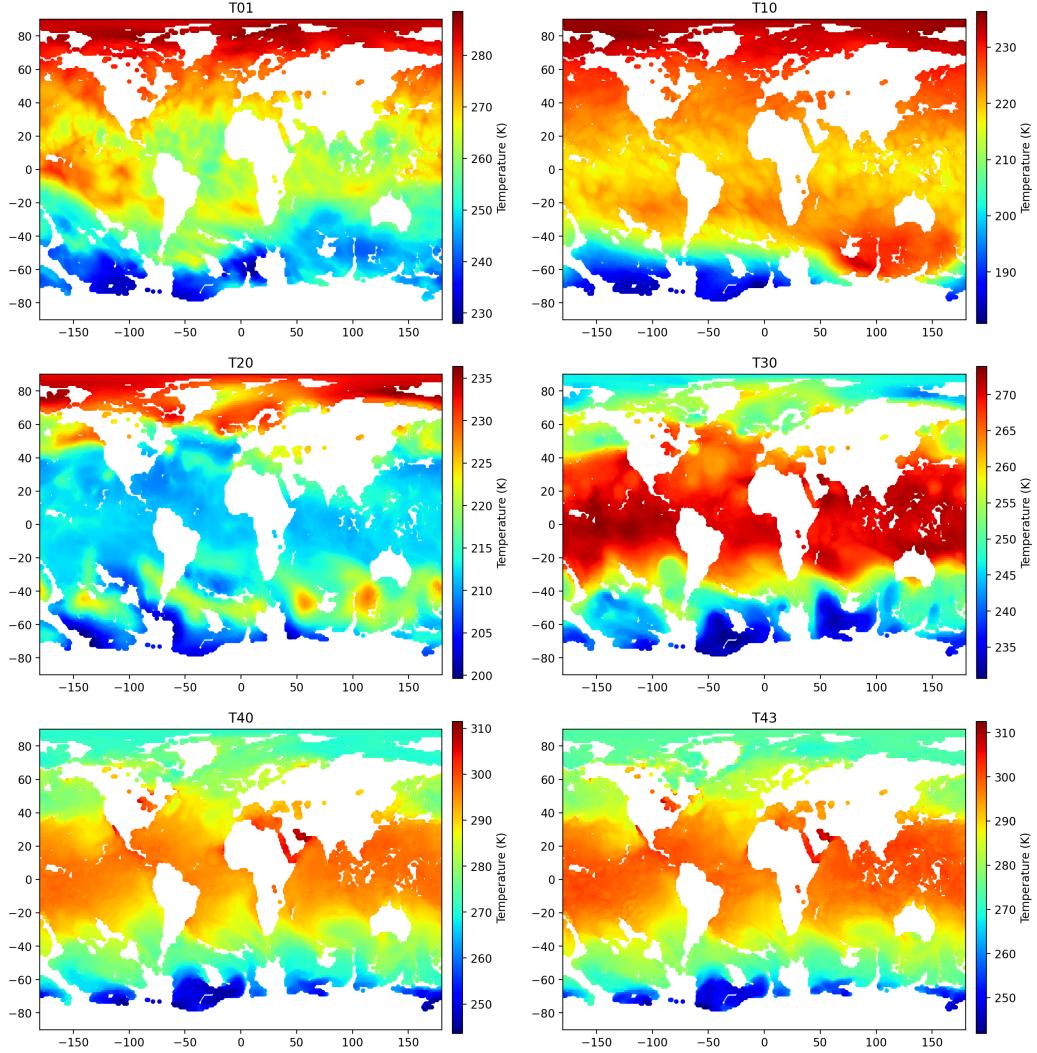


Figure 3: The temperature map of layer 1, 10, 20, 30, 40, 43.

troposphere (layer 20), this is the reason why clouds and rainfalls are formed in this layer. At higher altitude to the top of troposphere, the humidity concentrates along the equator, this phenomena is clearly consistent with the amount of precipitation recorded in these regions. At lower layer down to the surface, the humidity seem to be saturated in all the map. The shape of humidity profile was briefly discussed, one can also relate to the profile of temperature. The humidity starts to increase from 0 to its peak in stratosphere at the height where the temperature also change its pattern, this is seen more obviously in Figure 6 (right). We will analyse this correlation between humidity and temperature later on.

2.3 AMSU

Figure 7, 8, 9 shows the map of different window channels, oxygen and water vapour lines in the two instrument AMSU-A and AMSU-B. The temperature measured was derived from the total emission received in each channels. The window maps indicate the radiation coming mostly from surface and the boundary layer. Therefore the signals in these channels are quite weak, except for the equator regions, where we can see very strong signals in all window channels.

For AMSU-A, rom Figure 10 the height at the peak of weighting function decreases as

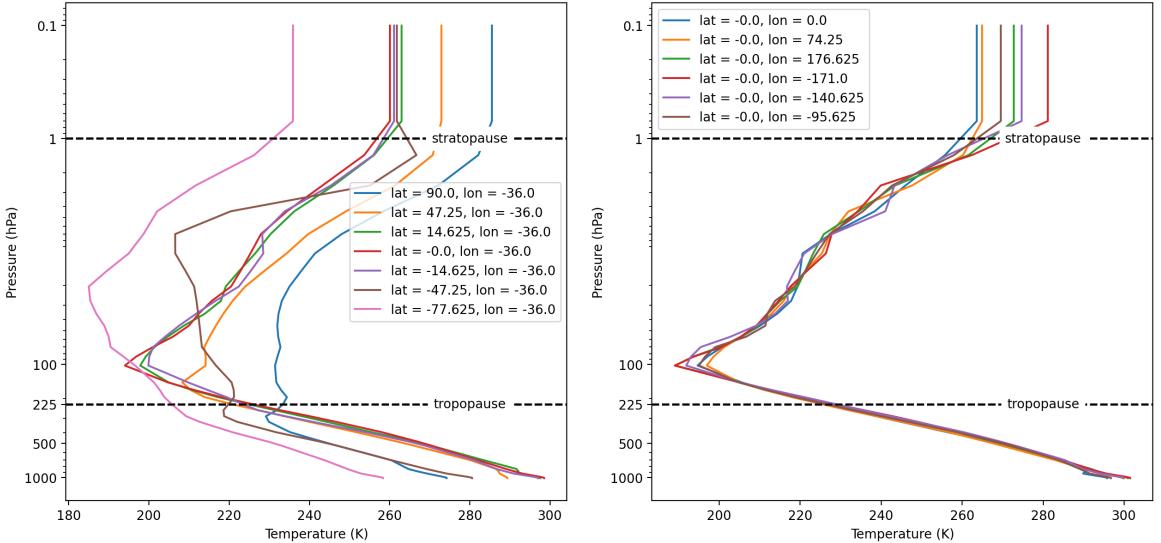


Figure 4: Vertical temperature profile up to 52km above the sea level of the atmosphere at the same longitudes, different latitudes (left) and different longitudes, same latitudes (right). The horizontal dash lines for stratopause and tropopause are just relative, not the exact positions.

frequency decreases. It implies that each oxygen channels show the radiation from oxygen molecules at different layers in the atmosphere. Channels 52.8, 54.4, 54.9, 55.5 GHz represent for the emission in troposphere and channels near 57.2 GHz represent for the stratosphere. While with AMSU-B, the measurements are coming from the troposphere where we have the appearance of water vapour. One can see some analogies between these maps and the maps of geophysical data (Section 2.1, 2.2). The spectrums of these instruments also have the same shapes at across different latitudes, the causes for the shifts in each spectrum was discussed in previous sections, mostly due to the illuminating area of the Sun.

The purposes of AMSU-A and AMSU-B are supplying a high quality data for retrieving the atmospheric temperature and moisture profile in real time, in which AMSU-A uses to deduce the temperature and AMSU-B responsible for humidity. In the next parts, we will see the relationship between the measurements from these instruments with the geophysical data and then trying to find a model that help us predict real value from satellite data.

3 Correlation

In order to get an idea for which data we should use to get our prediction, correlation matrix is a useful tools. This matrix show us the correlation between different set of measurements i.e. layers, channels,... to check whether the dependence between these measurements is strong or non-correlated. Figure 11 shows the correlation matrices between each channels and layers inside different datasets.

The temperature layers within the same atmospheric layers (troposphere, stratosphere, mesosphere) appear to be very correlated with each other with two bright corners (in top left plot). Also we can see the dark region with negative correlation indicate the opposite pattern between troposphere (including layers 20-43) and stratosphere(including layers 5-19). While almost the humidity layers doesn't have any relation (top right plot). The bright region can be neglected in this case, because it appears to channels, where don't have much water vapour concentration as discussed in Section 2.2. The correlation matrices of two

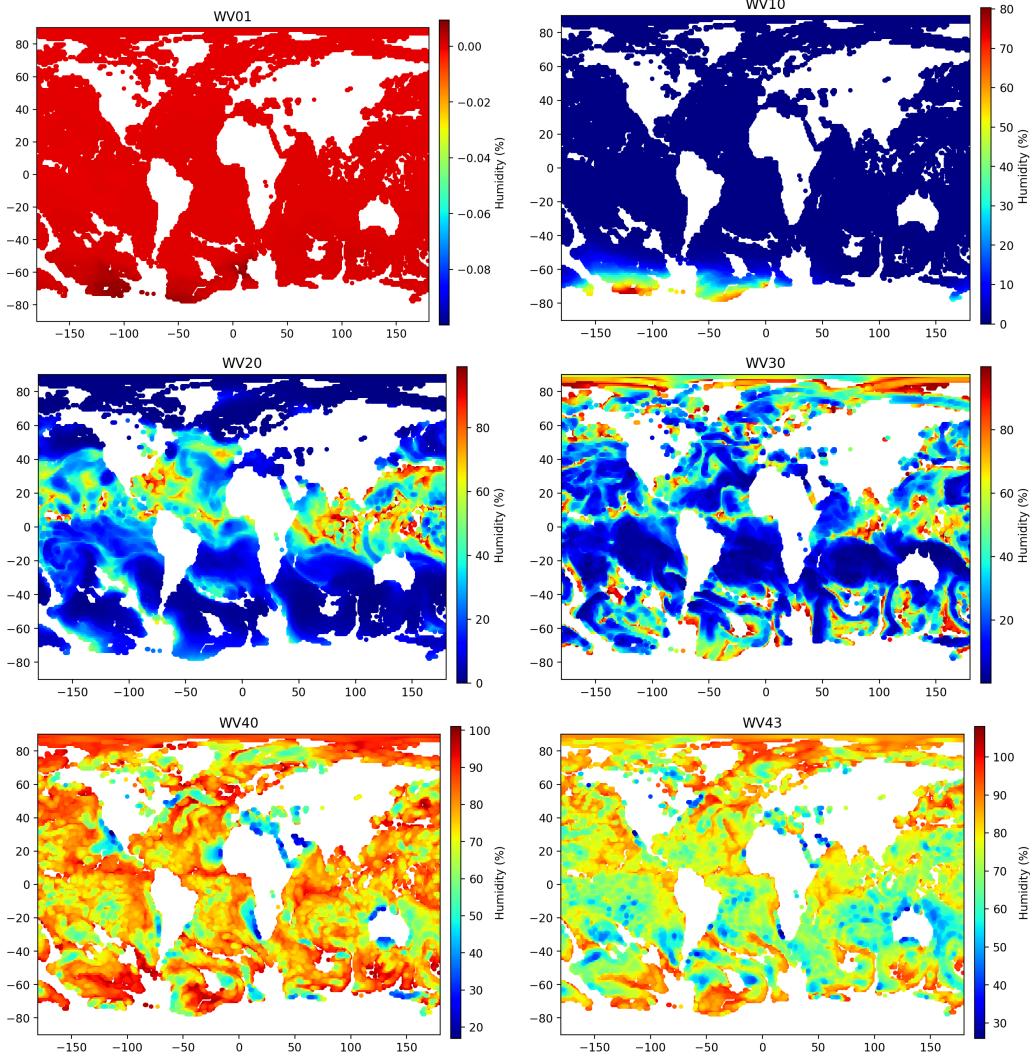


Figure 5: The humidity map of layer 1, 10, 20, 30, 40, 43.

instruments are pretty similar. Windows channels are very high correlated with itself and showing no relation with emission lines.

Figure 12 also represent the independent between temperature and humidity at the lower atmospheric layers to the surface, while strongly negative correlations are seen in the upper layers. Remind from Section 1.2, the window channels imply the radiation from the surface, it's clearly shown in the bottom left plot that the correlations between channels 1,2,3 and 15 with the lower layers are quite high. The plot also shows the highly dependence between oxygen radiation and atmospheric temperature (indicated by bright regions). In the other hand, it's unusual with AMSU-B since there is not much correlation between water vapour and humidity. The scatter plots in Figure 13 visualizes more clearly about the correlation of some pair samples.

4 Modelization

In this section, we explore the efficiency of different models use to predict the profile of temperature and humidity. In particular, the first approach is linear regression, a method that is pretty useful to get a linear model between parameters, which are highly correlated to each

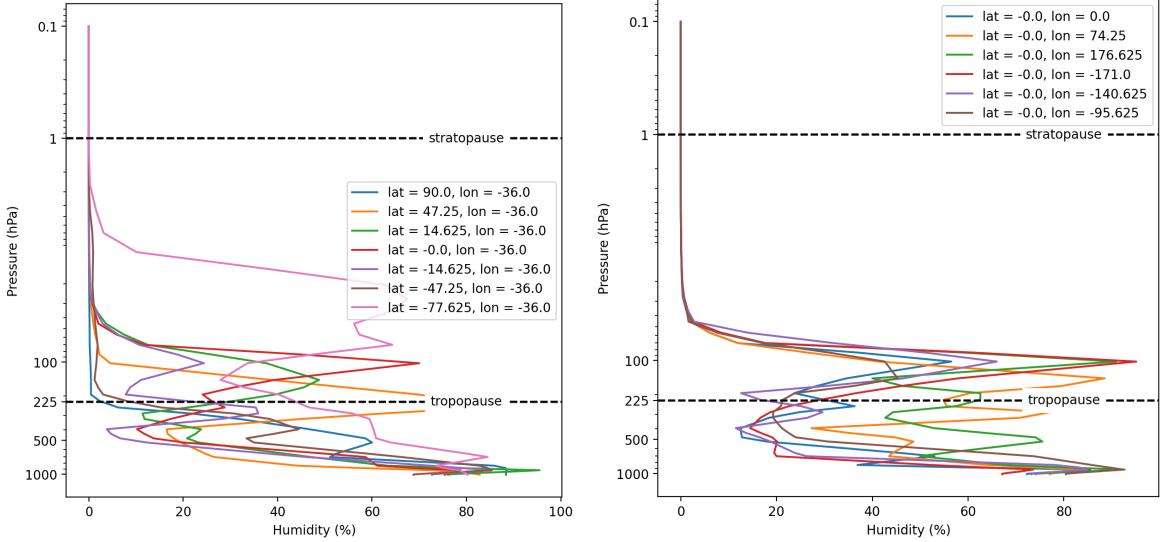


Figure 6: Vertical humidity profile up to 52km above the sea level of the atmosphere at the same longitudes, different latitudes (left) and different longitudes, same latitudes (right). The horizontal dash lines for stratopause and tropopause are just relative, not the exact positions.

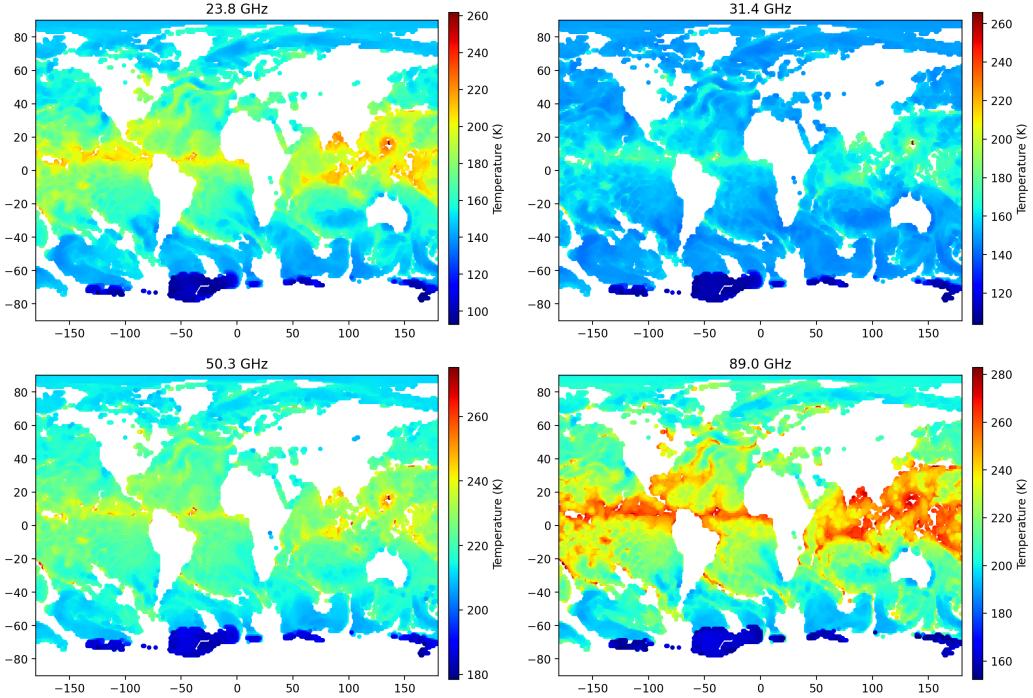


Figure 7: AMSU-A, four window channels at 23.8, 31.4, 50.3, and 89 GHz.

other. The second one is neural networks, a deep learning model simulates the human brain, consists of connected units or nodes called artificial neurons.

The linear regression uses input and output data to find the best set of parameters for equation 1, which produces the smallest error. In our case, the input value is the satellite measurements and we expect the output is equivalent to geophysical data.

$$model = w_1 * channel_1 + w_2 * channel_2 + w_3 * channel_3 + \dots \quad (1)$$

We use all the satellite channels to estimate the coefficients regardless their correlation

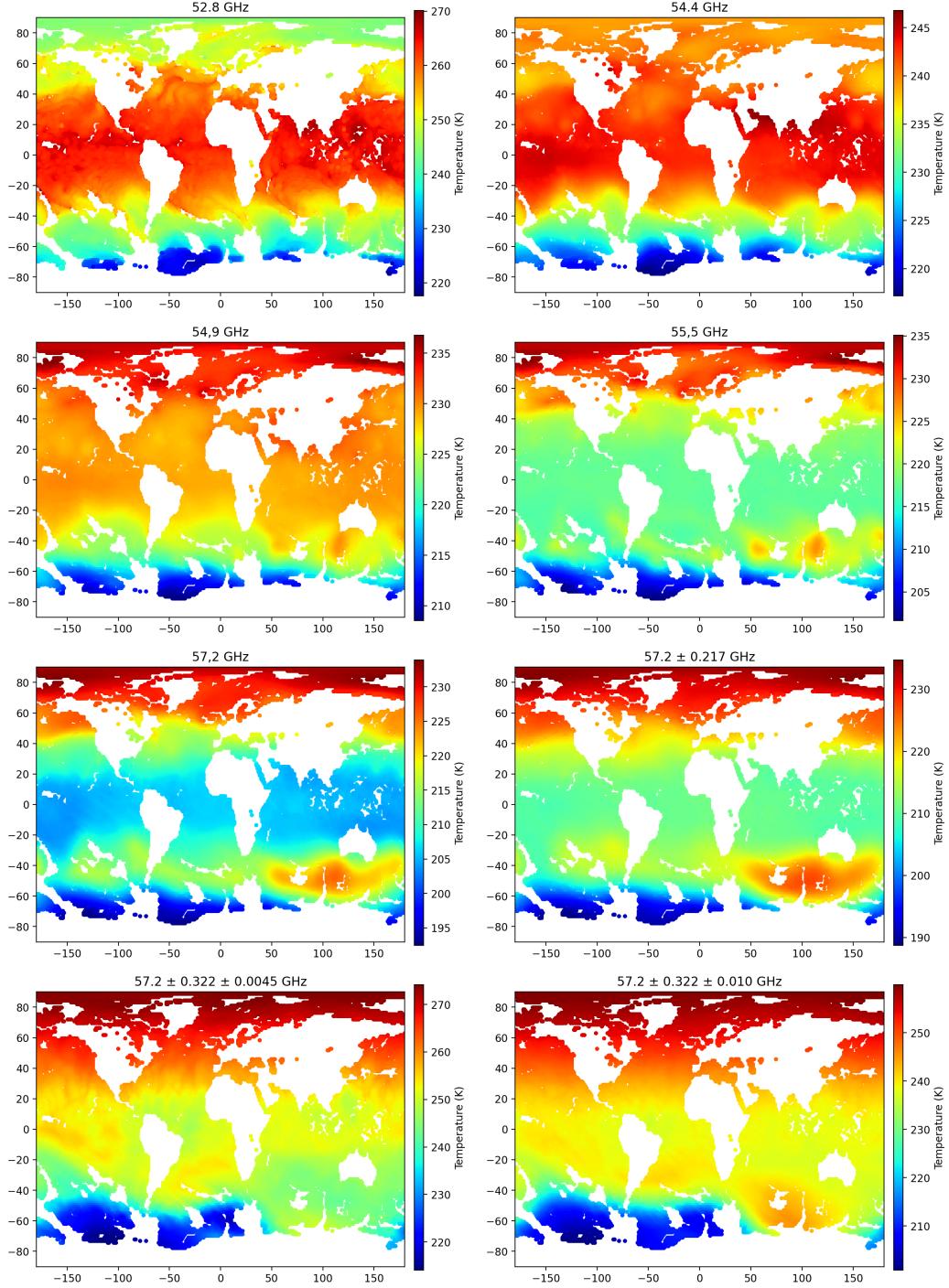


Figure 8: AMSU-A, 8 channels located close to the oxygen absorption lines below 60 GHz.

with the geophysical data as shown in previous section. The coefficients are represented in Figure 14. We find a huge range of frequency that has highly contribution to the prediction of temperature using AMSU-A instrument, especially for channel 6 and 7 corresponds to 54.4 and 54.9 GHz. The estimated errors also presented in Figure 15, the errors of temperatures derived from the model mostly range from -2 to 2 K, except for high latitude regions near the South Pole. While the performance of the AMSU-B used to estimate the humidity is not so good.

Neural networks are composed of simple elements operating in parallel. These elements

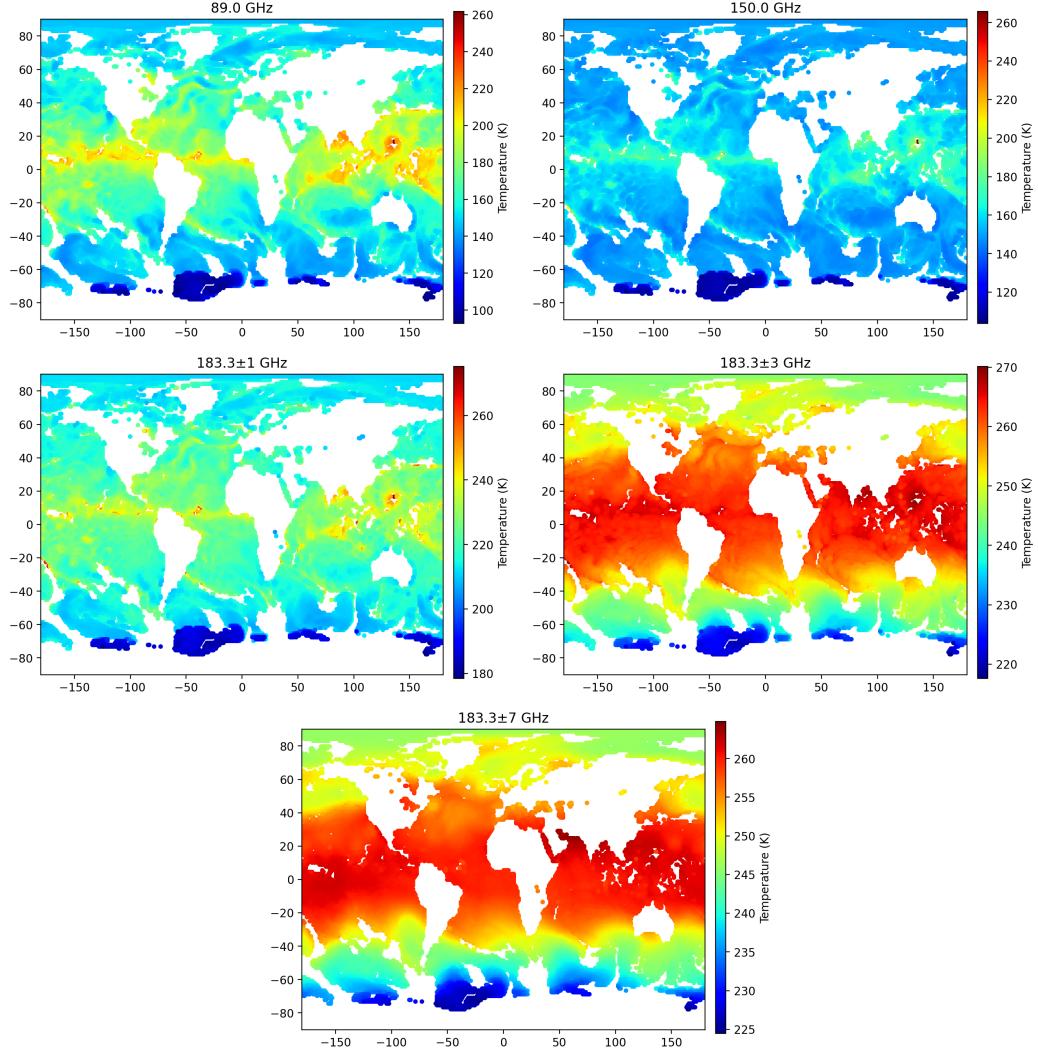


Figure 9: AMSU-B, two window channels at 89 and 150 GHz and three channels centered on the 183.31 GHz water vapor line.

are inspired by biological nervous systems. As in nature, the connections between elements largely determine the network function. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between elements. Typically, neural networks are adjusted, or trained, so that a particular input leads to a specific target output. The network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically, many such input/target pairs are needed to train a network. The comparison between linear regression and neural networks is shown in Figure 16. Linear regression yielded a better estimation for both temperature and humidity than neural networks.

5 Classification

Within this section, we use an unsupervised algorithm called K-means clustering to classify the Earth into different regions based on temperature and humidity profile. K-means clustering is a partitioning method. It partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. K-means clus-

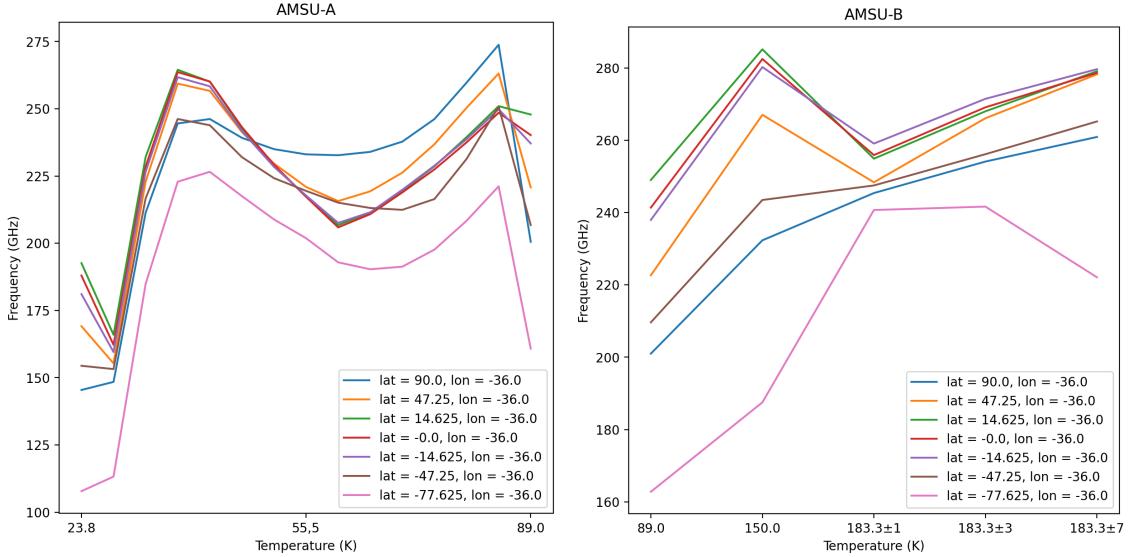


Figure 10: Spectral bands of AMSU-A and AMSU-B at different latitudes.

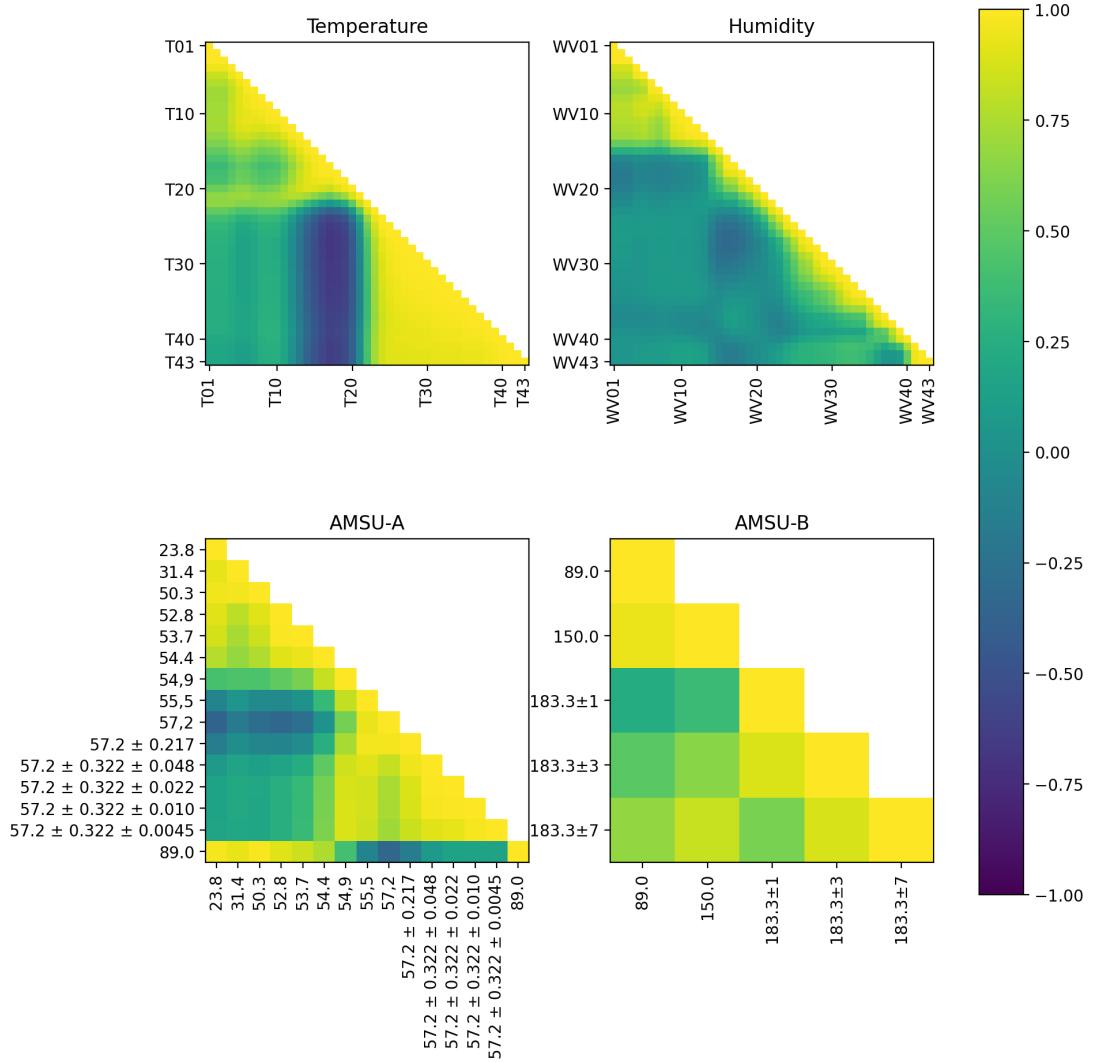


Figure 11: Correlation matrices between different layers and channels for each measurement.

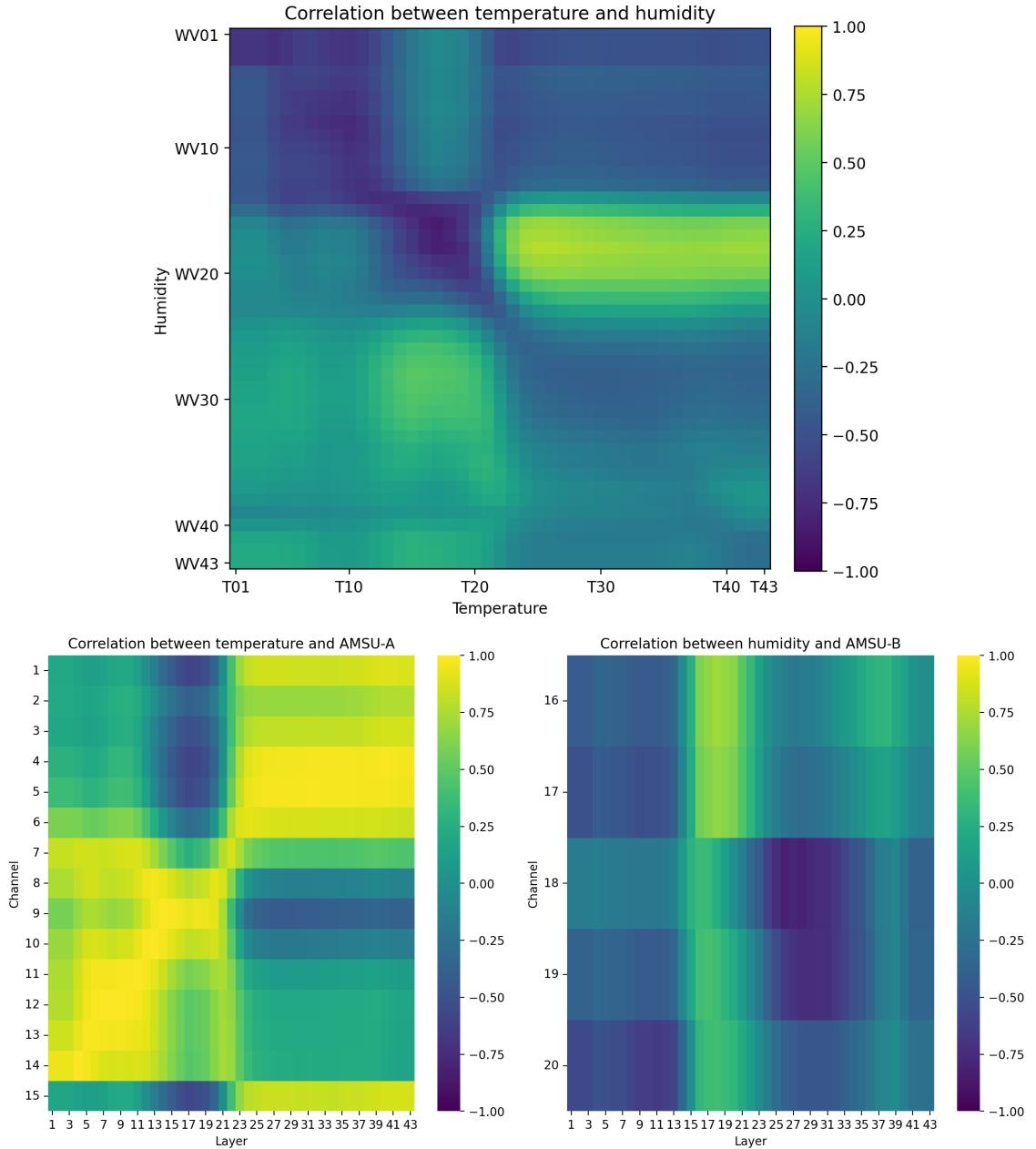


Figure 12: Correlation between geophysical data and satellite data.

tering treats each observation in the data as an object having a location in space. It finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. Cluster centroids are computed differently for each distance measure, to minimize the sum with respect to the specified measure.

The results are shown in Figure 17. Applying K-means to temperature data, we can basically divide the Earth into 5 distinguish regions along the latitudes. While with humidity data, there are some messy between these classes.

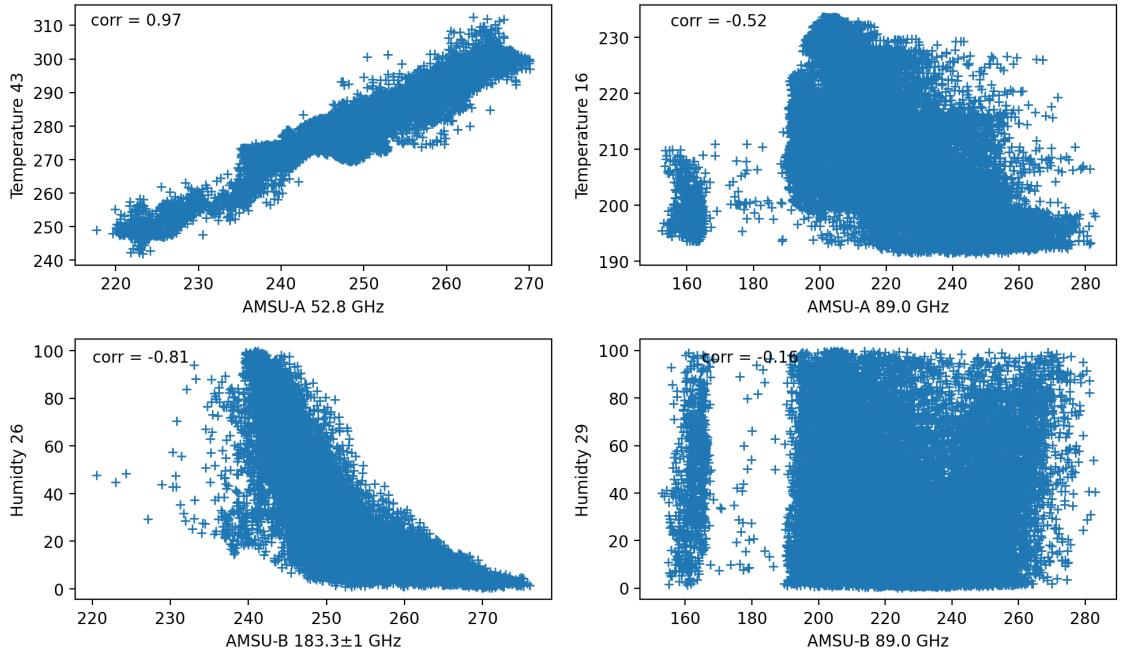


Figure 13: Scatter plot, top left: high correlation, top right: negative correlation, bottom left: negative correlation, bottom right: no correlation.

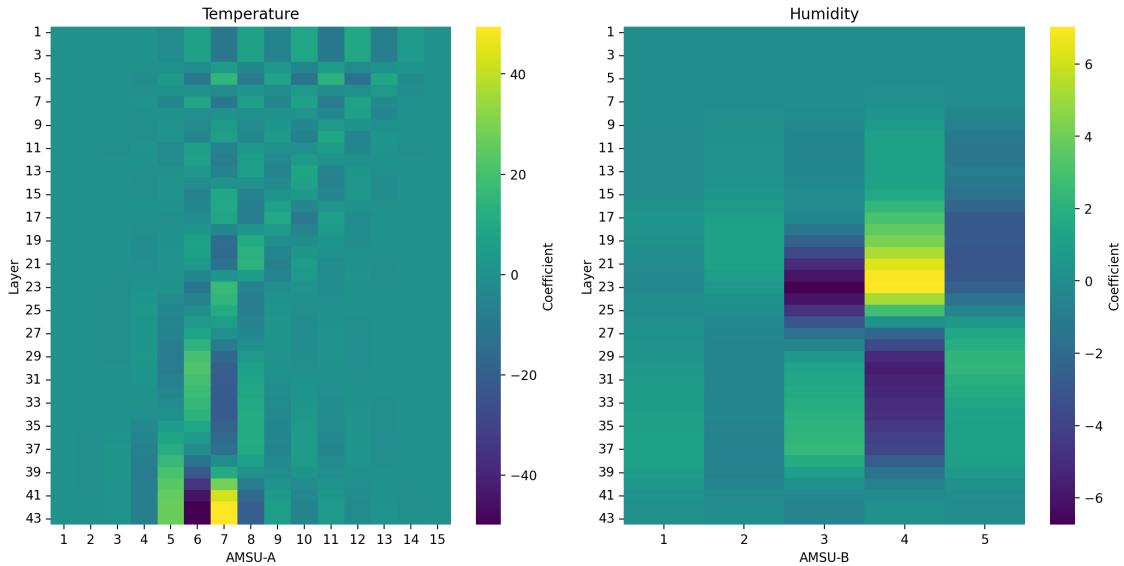


Figure 14: Coefficients distribution of linear model predicted by linear regression.

6 Principle Components Analysis

In data sets with many variables, groups of variables often move together. One reason for this is that more than one variable might be measuring the same driving principle governing the behavior of the system. In many systems there are only a few such driving forces. But an abundance of instrumentation enables to measure dozens of system variables. When this happens, this redundancy of information can be taken advantage. The problem can be simplified by replacing a group of variables with a single new variable. Principal component analysis is a quantitatively rigorous method for achieving this simplification. The method generates a new set of variables, called principal components. Each principal component is a

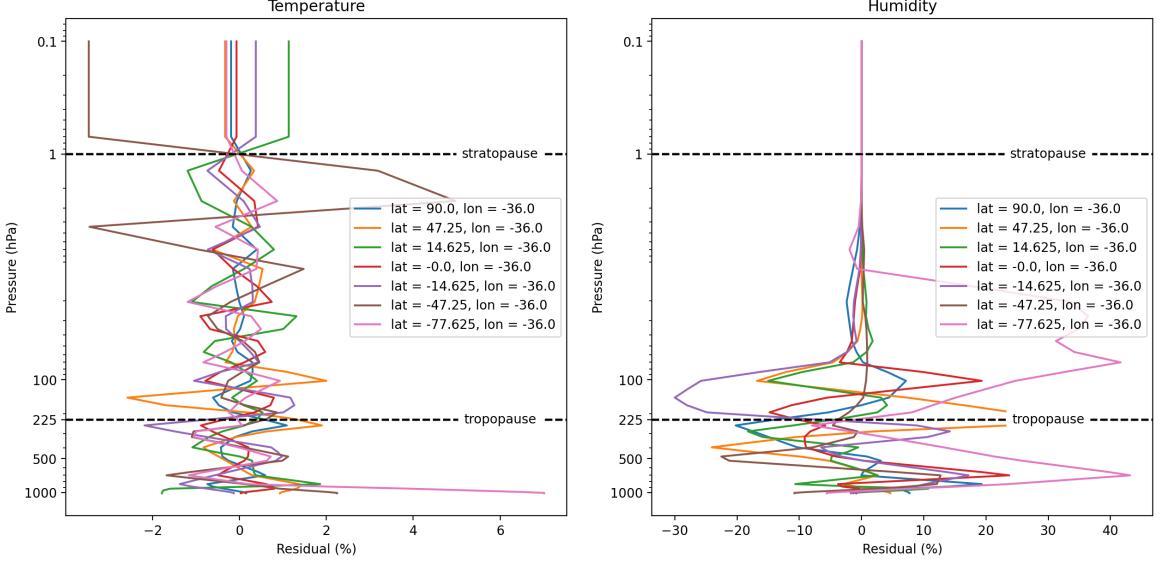


Figure 15: The errors between predicted value from linear regression and the geophysical value.

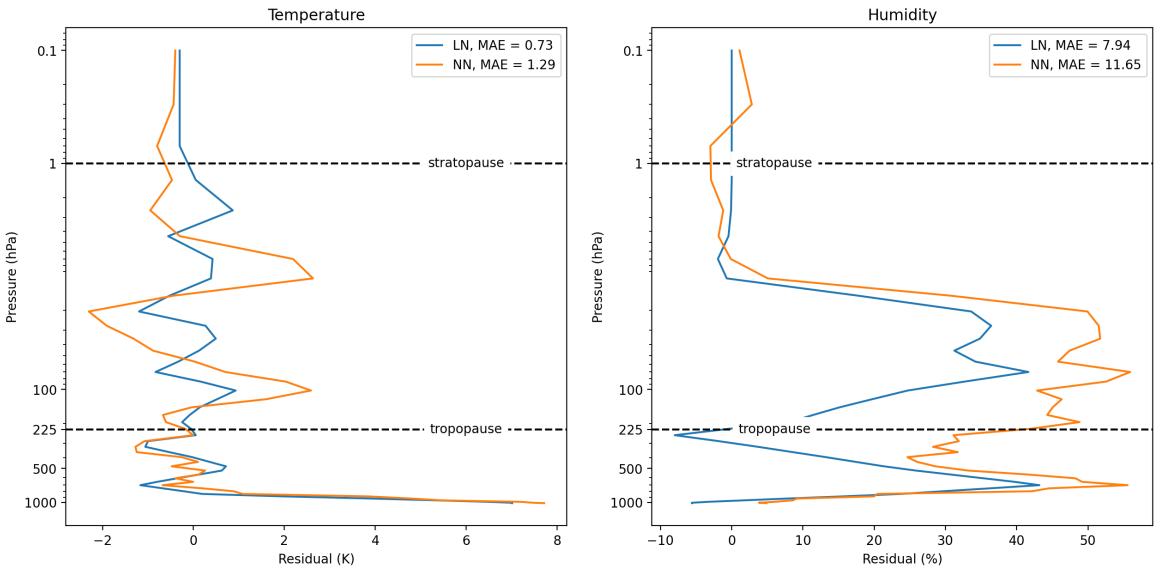


Figure 16: Comparing the errors predicted by linear regression and neural networks near the South Pole. The mean absolute value also presented.

linear combination of the original variables. All the principal components are orthogonal to each other, so there is no redundant information. The principal components as a whole form an orthogonal basis for the space of the data. The first principal component is a single axis in space. When each observation is projected on that axis, the resulting values form a new variable. And the variance of this variable is the maximum among all possible choices of the first axis. The second principal component is another axis in space, perpendicular to the first. Projecting the observations on this axis generates another new variable. The variance of this variable is the maximum among all possible choices of this second axis. The full set of principal components is as large as the original set of variables. But it is commonplace for the sum of the variances of the first few principal components to account for most of the total variance of the original data. Thus, the original set of variables can be represented by

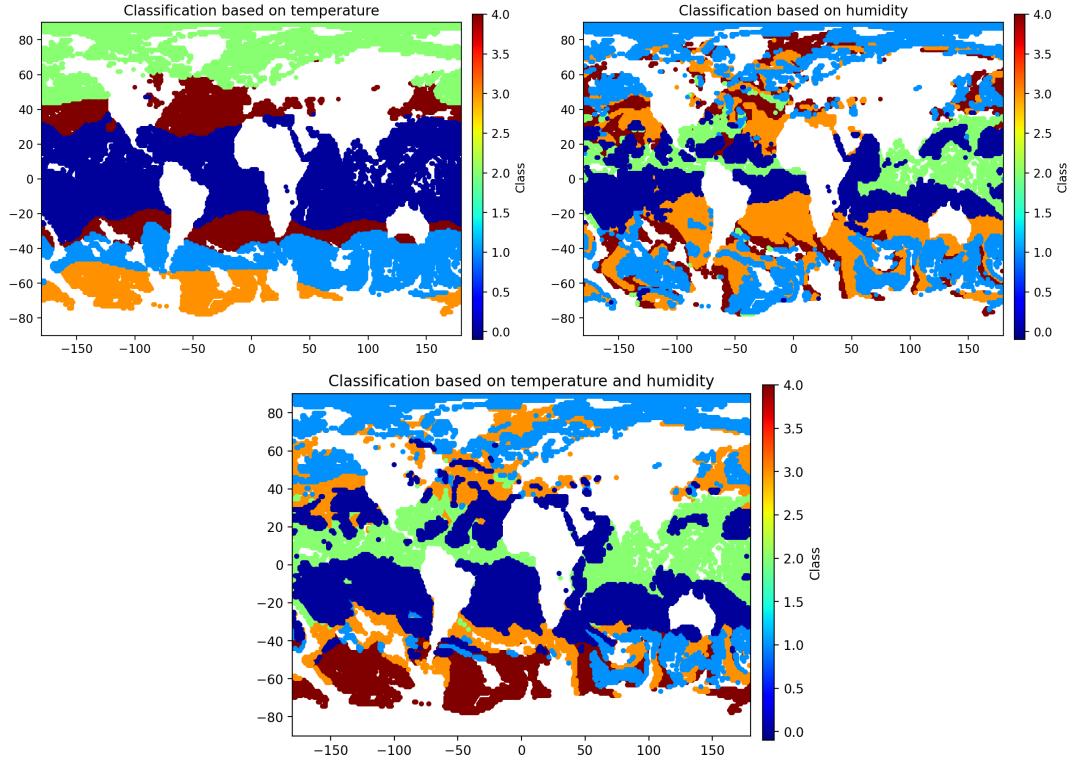


Figure 17: Running K-means clustering to temperature and humidity and both of them together.

a smaller set of principal components without losing much information. As a consequence, principal component analysis can be used for data compression. The results are shown in Figure 18, 19

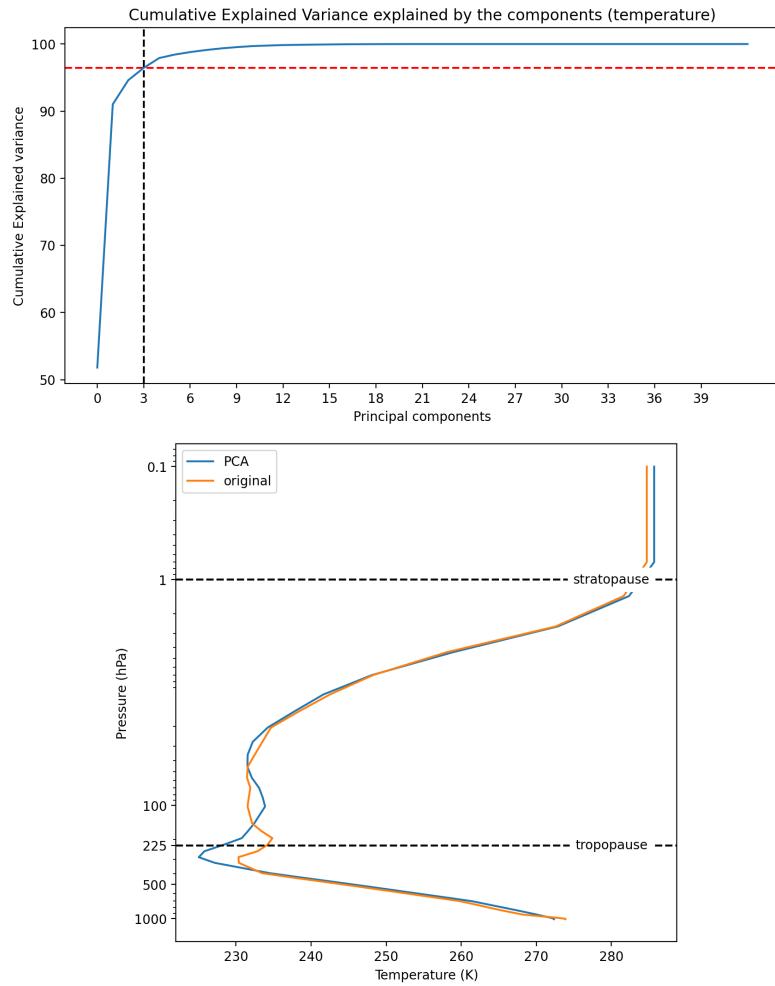


Figure 18: Running PCA for temperature.

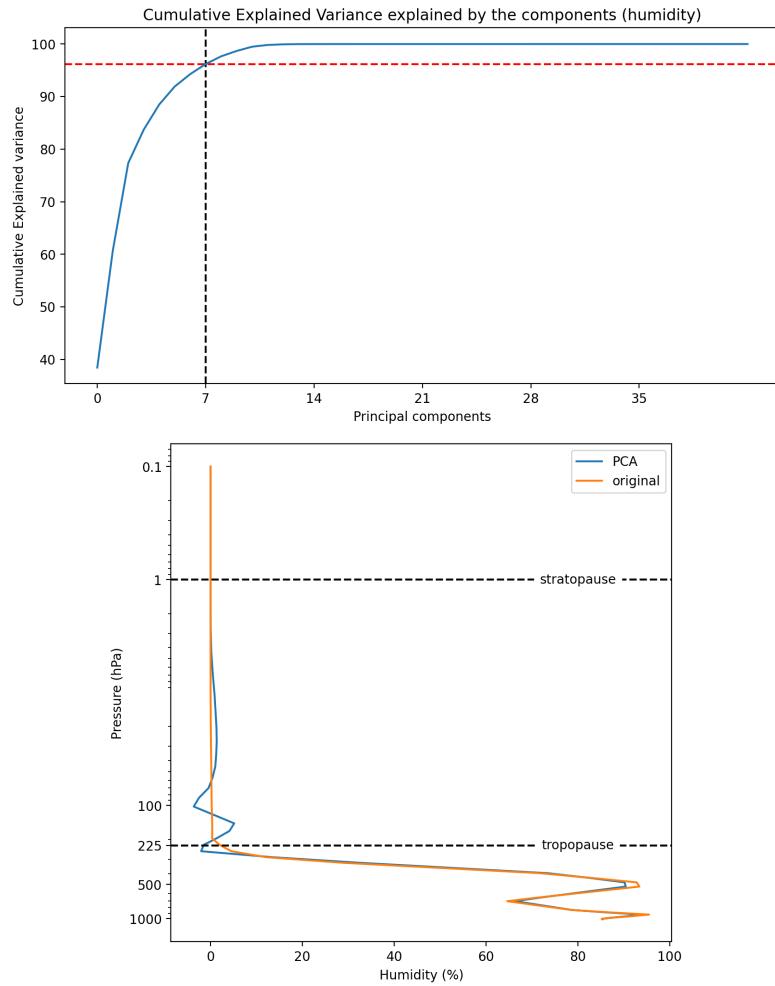


Figure 19: Running PCA for humidity.