Indian Institute of Technology, Bombay

Department of Electrical Engineering

# Wavelets in 1-D Convolutional Neural Network

## EE 678 Wavelets

Guide: Prof. Vikram Gadre

Supervisor: Kishore Tarafdar


| Team: | Roll Number: |
|---|---|
| Aditya Anavkar | 19D070004 |
| Puranjay Datta | 19D070048 |


ACADEMIC YEAR 2022-2023

ABSTRACT

**Ground-based Microwave Radiometer can provide a vertical profile of atmospheric variables such as relative humidity, pressure and temperature with very high time resolution (every couple of minutes). These measurements can greatly help determine climate conditions on a dynamic basis. In this research exposition, we have utilized wavelets in conjunction with neural networks to propose the model Modified Wavelet Convolutional Neural Network (MCWNN) to retrieve the vertical profile of the atmosphere. As opposed to earlier literature, our model is very lightweight with a sparse number of parameters. Validation of the model on test data gave percentage errors of 0.37%, 0.61% and 40.53% for the prediction of pressure, temperature and relative humidity respectively. We have also demonstrated the comparison of our model with other architectures and previous works to retrieve vertical profiles from MWR data**


## I. INTRODUCTION

The vertical profiles of atmosphere mainly temperature and relative humidity needs to be monitored to understand how climatic systems have evolved over time. Temperature and humidity structure are key inputs for numerical weather prediction models as they affect the stability of the atmosphere. In addition to their temperature and humidity profiles, knowledge of their time-based changes is particularly useful for studies of the atmospheric boundary layer.

Accurate atmospheric vertical temperature and humidity profiles are currently provided through the worldwide radiosonde network. Despite the excellent precision of the data derived from this source, the observations are limited by their expense and poor temporal resolution (typically once or twice a day), which is insufficient to capture the daily variation of the atmospheric profile. A more contemporary method that may be used to collect crucial information on weather profiles in the lower part of the atmosphere is a ground-based Microwave Radiometer (MWR). A MWR is a type of radiometer that measures the energy that is emitted at microwave frequencies, which range from 0.3 to 300 GHz. To determine the distinctive emission spectra of atmospheres, surfaces, or alien objects, they are often outfitted with numerous receiving channels.

Since there is no closed function that directly extracts the atmospheric profile from the MWR measurements, we must make an educated approximation. Several strategies have been tested in recent years to increase the accuracy of retrieval algorithms in MWR methodologies. For the model's nonlinear interactions, which are necessary for the recovery of humidity profiles, neural networks among these techniques can provide a greatest framework. In this research exposition, we investigated how to estimate the model using wavelets in combination with neural networks.

## II. LITERATURE SURVEY

### A. SAMEER

The team at Society for Applied Microwave Electronics Engineering and Research (SAMEER) is in the process of building the Microwave Radiometer for ground-based retrieval of the vertical profile of atmospheric features. There are two sides to the operation the Forward model and the Inverse model. The fundamental measurement obtained from MWR are brightness temperature $T_b$ at different microwave frequencies. The estimation of brightness temperatures at different Frequency bands from the Vertical Temperature and Humidity Profiles is known as Forward Model. The forward Model mainly comprises Radiative Transfer Equations and Weighting Functions. These weighting functions are dependent on Atmospheric Absorption and the Radiometric Channel Frequency. In the frequency range from 20 to 200 GHz, water vapour, oxygen, and cloud liquid are the main sources of atmospheric emission and absorption. A visual representation of the working of the Forward model is depicted below.

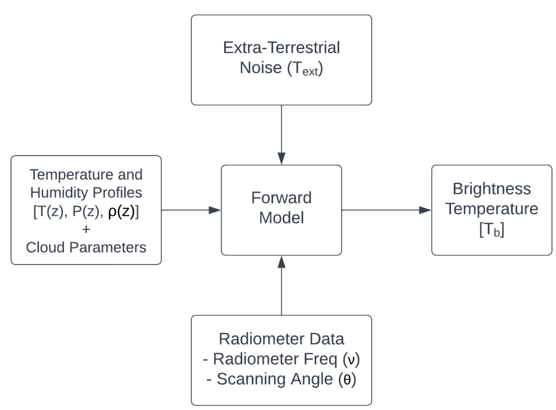



Figure 1: Forward Model,Source

Here, absorption as water vapour is used for humidity profiling we use the 20-30 GHz band which consists of 8 channels. The absorption due to oxygen is used for temperature profiling and consists of 14 channels in the range of 50-60 GHz. While the forward model holds importance as a measurement of the microwave radiation, we are more interested in the Inverse model of the Microwave Radiometer (MWR) which retrieves atmospheric profiles from measured brightness temperature.

On $14^{th}$ of October I, Aditya Anavkar visited the SAMEER lab along with Mr Kishore where we met the team working on the project. They walked us through the current progress of the project so far and their HOD Mr Anil Kulkarni gave an in-depth breakdown of the problem at hand which helped me understand various aspects of the project. Mr Tanmay has explored using Multivariate Linear Regression and ANN models for the retrieval of the temperature profile from the MWR data wherein they arrived at a final error of 3.3K. A major limitation of the model was that it was built only for a single height range for example 100m and a new model would be required for every new height. The task of building and improving models for retrieving pressure, relative humidity and temperature was explored in this exposition. The experience at SAMEER was gratifying and helped us learn a lot; all these learnings helped accelerate the research work.

### B. *Wavelet Pooling for Convolutional Neural Network*

*Travis Williams, Robert Li*

The categorization of images and objects rises continually with the introduction of Convolutional Neural Networks, but its steady use needs ongoing review and updating of fundamental ideas. Typically, convolutional layer operations are the focus of network regularisation algorithms, leaving pooling layer activities without viable alternatives. In contrast to average pooling or maximum pooling, the proposed research addresses the use of discrete wavelet transformations in combination with downsampling to provide arguably superior pooling output. Data sets like CIFAR[12] and MNIST[13] are used to test the concept. This approach reduces features in a more structurally compact manner than pooling via neighbouring regions, addressing the overfitting issue that max-pooling encounters.

The two most widely used strategies for pooling are average pooling and maximum pooling. In average pooling, an area is chosen for the condensed feature map based on its average value. Max pooling is the process of choosing the condensed feature map of the region R with the highest value. Max and average pooling both have drawbacks while being efficient, straightforward techniques. Depending on the data, max pooling may remove some visual features. This occurs when the important details are less intense than the minor ones. Max pooling frequently overfits training data as well. Depending on the data, average pooling may muddy up important features in the data.

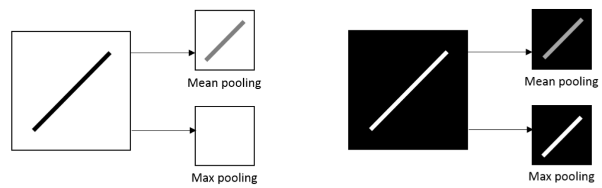



Figure 2: Shortcomings of Max and Average Pooling, Source

Wavelets are used in the suggested pooling approach to reduce the size of the feature maps. The wavelet transform is suggested by the authors as a way to reduce neighbourhood reduction artefacts. The method discards the first-order subbands to capture the data compression more naturally.

Their suggested solution performs within respectable ranges of the other pooling methods in the SHVN dataset, outperforms all other pooling methods in the MNIST dataset, and beats all but one of the pooling methods in the CIFAR-10 datasets. The inclusion of batch normalisation and dropout demonstrates how the suggested approaches respond to network regularisation. It surpasses all but one of the pooling algorithms used in the CIFAR-10 dataset's non-dropout scenarios and comes close to matching them in the SHVN dataset. Their findings are consistent with those of earlier research showing that no single pooling technique performs better than others depending on the dataset and network configuration. To improve efficiency, many networks also alternate between pooling methods. It may be possible to change the wavelet basis in this area in the future to investigate which basis performs the pooling the best. Better image feature reductions can be achieved by adjusting the up-sampling and down-sampling parameters in the analysis and synthesis.

### C. *Retrieval of Temperature and Relative Humidity Profiles from Microwave Radiometer*

*Xing Yan, Liang Chen, Nana Lou*

The Microwave Radiometer device offers several benefits

its continuous readings, which may offer high temporal resolution data (data with a precision of 1 min or less), and its ability to perform in most weather circumstances except for heavy rainy days or other severe climatic conditions. To recover temperature and humidity profiles using data from a ground-based MWR, the authors of this study have suggested a deep learning method termed "BRNN". BRNN strives to reduce overfitting and offers a larger capacity to characterise nonlinear correlations between MWR measurements and atmospheric structure information than the conventional backpropagation neural network, which has previously been used for MWR profile retrieval. Four layers make up the created neural network: an input layer, two hidden layers, and an output layer. The gathered information, which consists of 17 features, is delivered to the input layer. The brightness temperature $T_b$ data from the MWR's 14 channels make up 14 of these features, while the surface pressure, temperature, and RH make up three of them.

With a root-mean-square error of 1.70 K for temperature and 11.72 for relative humidity, the radiosonde validation of BRNN reveals a good retrieval capacity. Using the BRNN approach, the temperature's RMSE was measured to be around 1 K below 2 km and less than 2 K below 8 km, although it had a higher RMSE in the upper atmosphere (above 3.5 km). At 10 km, the RMSE of BRNN was discovered to be 3.5 K. The RH bias of the BRNN approach is constant below 3 km and reaches up to 10 km. The article demonstrated that BRNN is an effective method to retrieve the temperature and humidity profiles.

## III. Problem Statement

During the course project/ SRE we have been closely working with SAMEER who is building the MWR device and Mr Kishore. There are three atmospheric variables namely pressure, temperature and relative humidity let them be denoted by $x_1$, $x_2$, $x_3$ and collectively denoted by vector $\vec{x}$. The microwave radiometer has 22 channels starting from 22.234 GHz to 58.800 GHz let them be denoted by $y_1$, $y_2$, ... $y_{22}$ and collectively denoted by vector $\vec{y}$. The Forward Model $\vec{y}$ = $T_{forward}$ ($\vec{x}$) is well established and backed by physical laws i.e. given the atmospheric variables we can determine the microwave readings. The inverse model on the other hand $\vec{x}$ = $T_{inverse}$ ($\vec{y}$), deriving atmospheric condition given the microwave radiometer reading is not straightforward and thus we need to try to approximate it with machine learning methods. This can be formulated as a multi-dimensional vector to vector mapping where we are mapping a vector of $T_b$ from the MWR channels to the vector of actual pressure, temperature and relative humidity. The main problem encountered with using MWR for vertical profiling is that it is a listening device i.e. it listens to the radiation coming from the atmosphere but has no explicit information about which height it came from. Thus by some alternate method, we also need to somehow incorporate height while building our model. In the last few years, multiple approaches have been tried to improve retrieval accuracy and methods for MWR. For modelling nonlinear relationships, which are crucial for retrieving vertical atmospheric profiles, neural networks among these techniques are found to perform the best.

## IV. Universal Approximation Theorem

In layman's terms, a single hidden layer is sufficient to model any continuous function with epsilon error under the supremum norm. Let $\sigma$ be any continuous sigmoidal function. Then finite sums of the form

$$G(x) = \sum_{j=1}^{N} \alpha_j \sigma(w_j x + \theta_j)$$

are dense in $C(I_n)$.In other words given any $f \in C(I_n), \epsilon > 0, \exists$ a sum $G(x)$ for which,

$$|G(x) - f(x)| < \epsilon \;\; \forall x \in I_n, \;\; I_n = [0,1]^n$$

Where, $C(I_n)$ : space of continuous function on $I_n$.
A function $\sigma$ is sigmoidal if

$$\sigma(x) = 1 \; as \; x \to \infty$$

$$= 0 \; as \; x \to -\infty$$

$$\exists N, \exists w_j, \theta_j, \alpha_j \text{ for } j = 1, 2, ..N \text{ s.t } |G(x) - f(x)| < \epsilon$$

### A. Proof

Suppose we let $w_j \to \infty$ for $j = 1, 2....N$ then

$$\lim_{w_j \to \infty} \sigma(w_j x) = 0 \;\; for \;\; x \leq 0$$

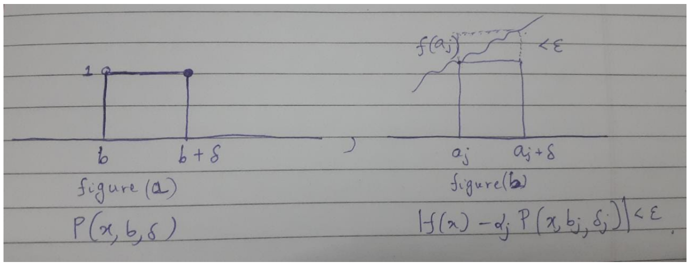
$$= 1 \;\; for \;\; x > 0$$

$$\lim_{w_j \to \infty} \sigma(w_j(x - b_j)) = 0 \; for \; x \leq b_j$$

$$= 1 \;\; for \;\; x > b_j$$

$$H(x) = \lim_{w \to \infty} \sigma(wx)$$

$$Define \;\; H(x, b) = \lim_{w \to \infty} \sigma(w(x - b))$$




We can use two such functions to create a piece

$$P(x, b, \delta) = H(x, b) - H(x, b + \delta)$$

Since $f(x)$ is continuous $\lim_{x \to a} f(x) = f(a) \;\; \forall a \in I_n$

$\exists$ an interval $(a_j, a_j + \Delta x)$ s.t

$$|f(x) - f(a_j)| < \epsilon \;\; \forall x \in (a_j, a_j + \Delta x)$$

choose $b_j = a_j, \delta_j = \Delta x, \alpha_j = f(a_j)$

$$|f(x) - f(a_j)| < \epsilon \;\; \forall \; a_j \leq x \leq a_j + \delta_j$$

$$|f(x) - \alpha_j P(x, b_j, \delta_j)| < \epsilon$$

Repeat the process for $x = a_{j+1} = b_j + \delta_j$.
Construct $G(x) = \sum_{j=1}^{N} \alpha_j P(x_j, b_j, \delta_j)$. Hence as the number of neurons increases exponentially ,$\epsilon$ decreases.

## V. Data and Methods

### A. Details of Dataset

The acquired data from SAMEER Lab consisted of two files, one meant for training the model and the other for testing, both of them are identical with respect to features. The data was recorded for 93 days; 22 MWR channels corresponding to $T_b$ readings, pressure, temperature and relative humidity readings from radiosonde for 150 vertical heights for each day. The data was split in two parts, training dataset which had 9389 rows and the testing dataset consisted of 4350 rows. The initial dataset had a lot of missing values and human error and thus required cleaning and pre-processing before it could be used to train the model. The features consist of height, followed by 22 channel brightness temperature readings and then Pressure, Temperature and Relative Humidity for that height. The procedure for creating the dataset was as follows -
The radiosonde device was allowed to move upward in the atmosphere and when it reached height $r$, it recorded the Pressure, Temperature and Relative Humidity readings. At the same time the Microwave Radiometer (MWR) recorded the $T_b$ readings for 22 channels. Thus we would have 22 channels and corresponding 3 atmospheric variables for the height $r$. This was then repeated for 150 heights every day. To get a better understanding

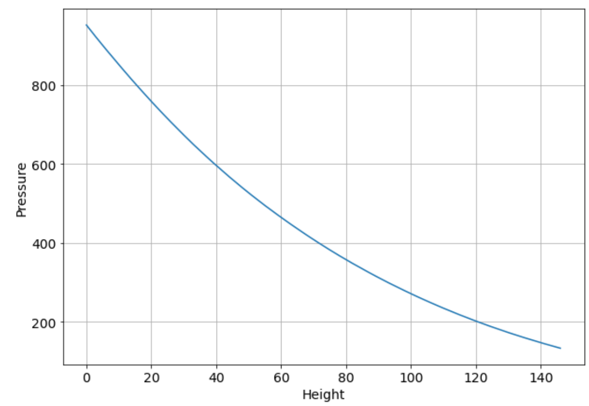



Figure 3: Pressure vs Atmospheric Height

of the data at hand, we plotted the weather features with height for a single day, with the height going from 0 to 150 units.
The pressure graph shows a somewhat exponential decrease with an increase in height which fits with the known pressure models. The temperature graph also shows a decrease with an increase in height. We can observe the schematic to show a very much linear pattern which makes it an ideal candidate to be fit with a Multivariate Linear Regression model

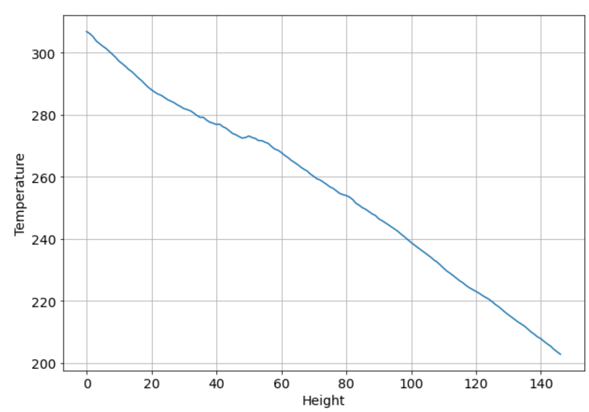



Figure 4: Temperature vs Atmospheric Height

The Relative Humidity in Figure 5 in the contrast to the above two displays high volatility and variance, thus posing difficulties while modelling.

### B. Proposed Methodology

In the initial duration of the exposition, the proposed hypothesis was to build a model to predict the ver-

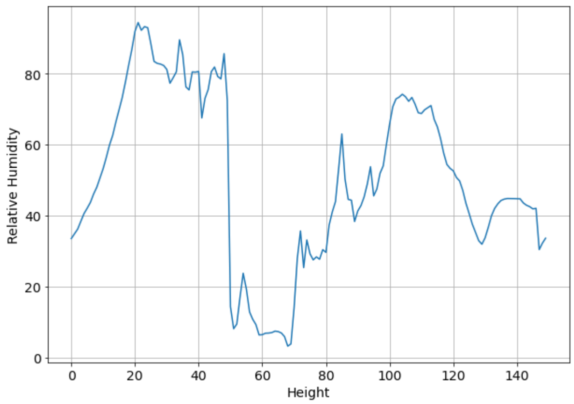



Figure 5: Relative Humidity vs Atmospheric Height

tical atmospheric profile using only the MWR channels following the empirical form $M(\vec{y}) = [p, t, rh]$ as discussed in the problem statement. We read literature regarding similar works and tried to build a model on the hypothesis but the network failed to capture the data appropriately. Later we proposed a slight modification adding height as a parameter into the model during the training phase, thus making the input a 23-length vector comprising of height followed by 22 brightness temperature from corresponding frequency channels of the MWR. The model can be viewed mathematically in the following line, where M denotes our model and p,t,rh denote pressure, temperature and relative humidity respectively.

$$M(h, y_1, y_2, ..., y_{22}) = [p, t, rh] \tag{1}$$

This change in feature selection gave a better result than the first model but was largely underfitting the relative humidity due to its high volatility. Following the suggestion by Mr Kishore that all the channels in the MWR are not equally useful for all variables, i.e. the frequency channels from 20 to 30 GHz (K band) carry information more relevant to relative humidity and frequency 50 to 60 GHz (V band) is more significant for temperature; we changed the training features to accommodate only the corresponding bands for the variables and changed the network to predict variables one at a time making it a vector to scalar mapping. Along with these, we introduced level 2 wavelet decomposition and LSTM layers in the network to get the final results.
We are interested in finding a function-to-function mapping for our problem, in particular. We will be trying to approximate the function by using Neural networks in conjugation with wavelets. In this exposition, we have proposed the Modified Convolutional Wavelet NeuralNet model (MCWNN) to retrieve the vertical atmospheric profile. In addition to our proposed MCWNN method, described earlier, we used several other machine-learning models for the comparison such as the multivariate linear regression model and a convolutional neural network with various activation functions were also implemented to compare results with.

### *Multivariate Linear Regression*

In this method, we are trying to find a linear correlation between atmospheric outputs and inputs namely 22 channels and height. Below is the generalized equation for the multivariate regression model

$$y = \beta_0 + \beta_1.x_1 + \beta_2.x_2 + \ldots.. + \beta_n.x_n \tag{2}$$

Where n represents the number of independent variables, $\beta_0$ to $\beta_n$ represents the coefficients, and $x_1$ to $x_n$ are the independent variable. In our case, n will be equal to 23.

### *Multilayer Perceptron Neural Network*

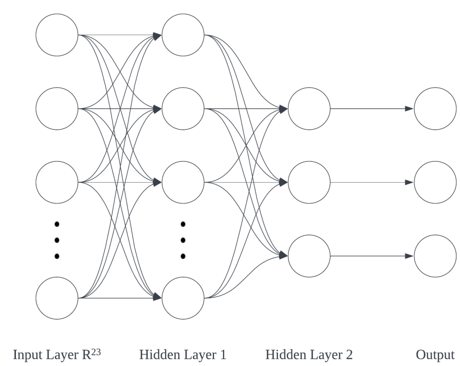



Figure 6: Multiperceptron Neural Network

The model comprises 2 hidden layers with 23 neurons in the first layer and 3 neurons in the second layer which has 552 +72 = 624 trainable parameters respectively. We tried and tested the model with various activation functions such as relu, sigmoid, mish and wavelets like morlet and mexican hat(admissible wavelets).
Like the universal approximation theorem which states that sigmoid function can model any continuous function we can use an admissible wavelet function which can implement both CWT and inverse CWT but might require more parameters to represent the function.

### *Wavelet Convolutional Neural Network*

This comprises of Level-1 decomposition where we train the corresponding scaling and wavelet branch independently and separately using a CNN which has kernel size = 30 to keep the number of training parameters equal to 624 for fair comparison, with batch normalization and Mish as activation function (which performed best amongst the deep-learning models). Used a flatten followed by a dense layer(3 neurons) in conjunction to get back the 3 outputs.

Adding the height parameter to the training data was crucial since the major weightage in predicting the output was of the height parameter(amongst the other 22 channels). This proved true when we analyzed the weights of the linear regression output for all the 3 atmospheric outputs.

We tried to learn the subbands independently which gave inferior results as compared to when we learnt all the sub-bands serially i.e by concatenating the previously learnt sub-band with the current sub-band. The ideology behind using wavelet decomposition was to extract the frequency spectrum-based relationship since the wavelet band corresponds to a high-frequency spectrum and the scaling function corresponds to a low-frequency spectrum. Learning the sub-bands serially has an added advantage in that a higher frequency spectrum component may be correlated and dependent on a lower frequency component hence might give us a better fit. Another logic for using a DWT decomposition is the resemblance of the downsampling to the pooling layer which is used for dimension reduction and enables the model to learn a particular subspace in a more abstract manner.

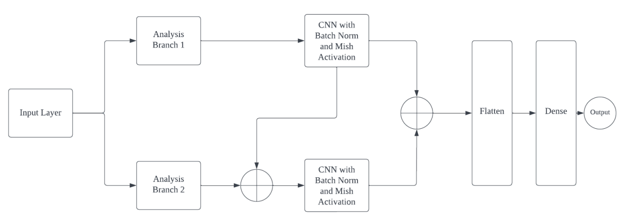



Figure 7: Wavelet Convolutional Neural Network

### *Modified Convolutional Wavelet Neural Network*

We modified the initial wavelet network to include Level-2 decomposition + LSTM + k-band,v-band training. LSTMs are used in place of CNN layers with the idea that they would better fit the high variance data of relative humidity. CNN was underfitting the data and was not able to capture the shape appropriately.

One of the reasons for substituting the CNNs with LSTM in this model was to make the model noise-resistant since LSTMs are like recurrent neural networks which keep the information of the past inputs (similar to a memory model). After sufficient training, we found that relative humidity which has a high variance had a better fit. Since wavelets are locally restricted in time, the LSTMs can serve as a good complement in terms of storing the temporal contextual information.

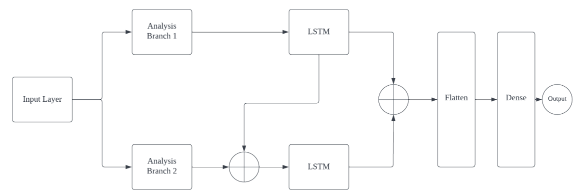



Figure 8: Modified Wavelet Convolutional Neural Network with LSTMs

## VI. Results and Discussion

The retrieval performance of all the models is compared in Table I for pressure, temperature and relative humidity. For each aforementioned model, the Mean Absolute Error (MAE) and Percentage Error (PE) for predictions on the validation data is illustrated. Our proposed MCWNN achieves the best accuracy in both MAE as well as PE when compared to all the models across all atmospheric parameters. It especially outperforms the rest in pressure wherein it achieves an absolute error of only 1.275 and a percentage error of 0.37%. It is noticeable for linear regression to perform well for predicting temperature with a percentage error of 0.9%

| | Method | MAE | PE |
|---|---|---|---|
| Pressure | I | 39.497 | 0.1231 |
| | II | 38.558 | 0.1180 |
| | III | 15.232 | 0.0401 |
| | IV | **1.275** | **0.0037** |
| Temperature | I | 2.262 | 0.0091 |
| | II | 2.590 | 0.0105 |
| | III | 6.950 | 0.0296 |
| | IV | **1.578** | **0.0061** |
| Relative Humidity | I | 14.365 | 0.4949 |
| | II | 14.553 | 0.5467 |
| | III | 15.098 | 0.5653 |
| | IV | **11.063** | **0.4053** |

Table I: Comparison of different Models
The models are I: Linear Regression, II: Multilayer Perceptron Neural Network, III: Wavelet Neural Network, IV: Modified Wavelet Neural Network

as was hypothesised earlier but performs unsatisfactorily for other parameters indicating it is not a suitable fit for them. Due to the high volatility of the relative humidity all the models have a high bias in predictions. The graphs for actual values versus predicted values for each weather parameter are illustrated ahead.

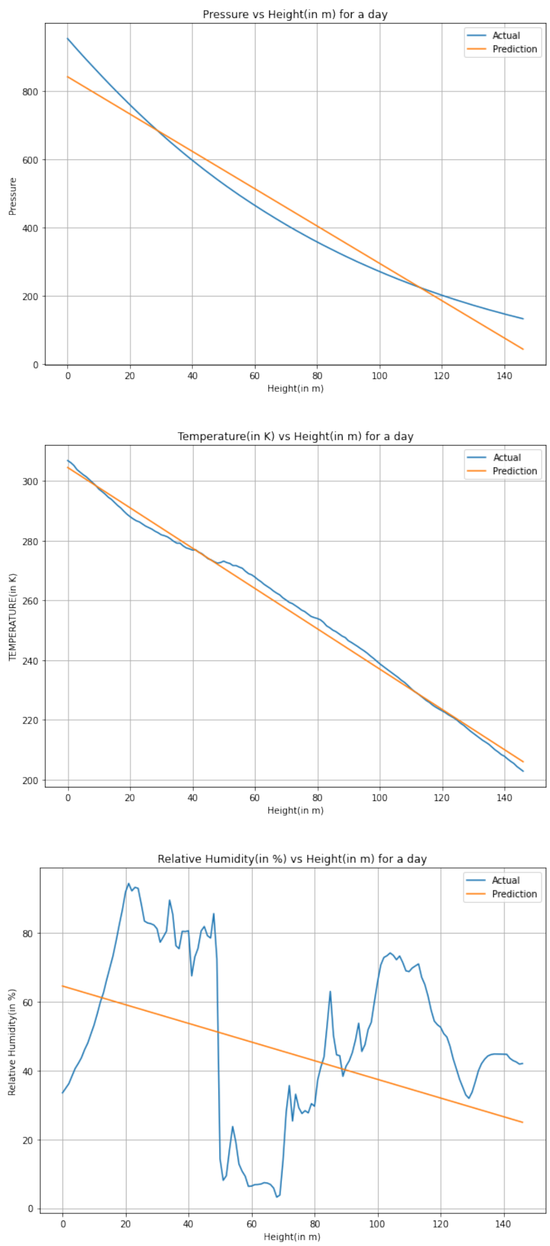



Figure 9: Linear Regression Model Graphs

Figure 9 shows the linear regression model to fit satisfactorily to temperature but cannot adapt to pressure or relative humidity. On similar lines we see the Multi-perceptron Neural Network to be under-fitting the atmospheric variables. This can be partially attributed to the network not being complex enough to learn the data due to sparse number of trainable parameters.

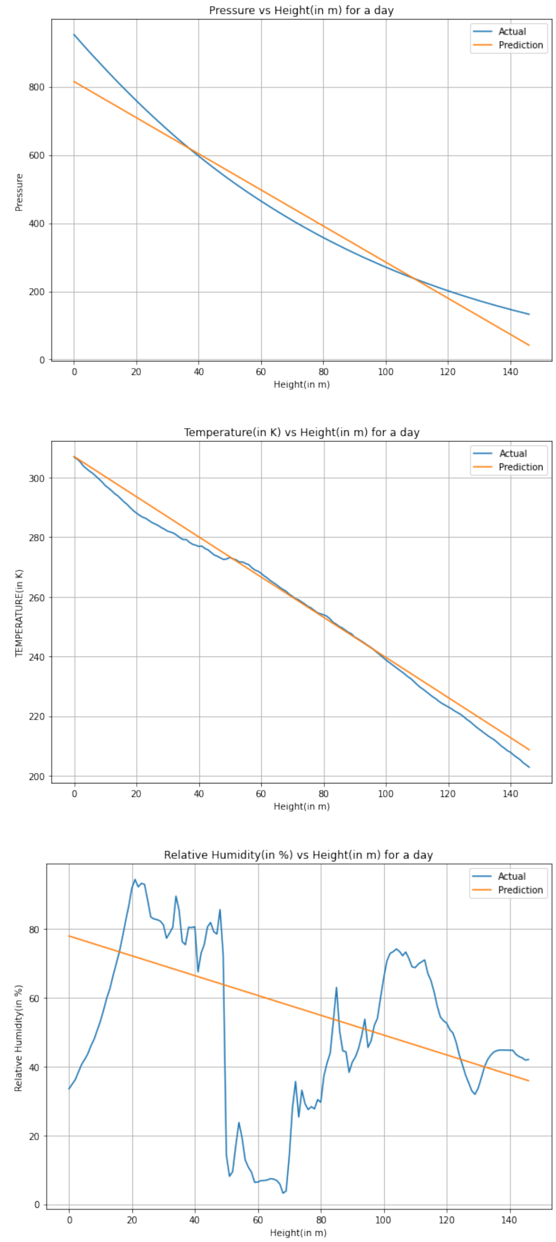



Figure 10: Multiperceptron Neural Network Graphs

The initial Wavelet Neural Network performed better in terms of pressure as compared to earlier models but suffered in temperature. Thus we decided to make changes in the architecture of the model. Finally, we have the plots for the proposed model MWNN in Figure 12. As we saw earlier it outperformed all the other models and we can see the same in the graphs. An almost perfect fit can be observed in the pressure graph between actual and predicted values. The addition of second-level

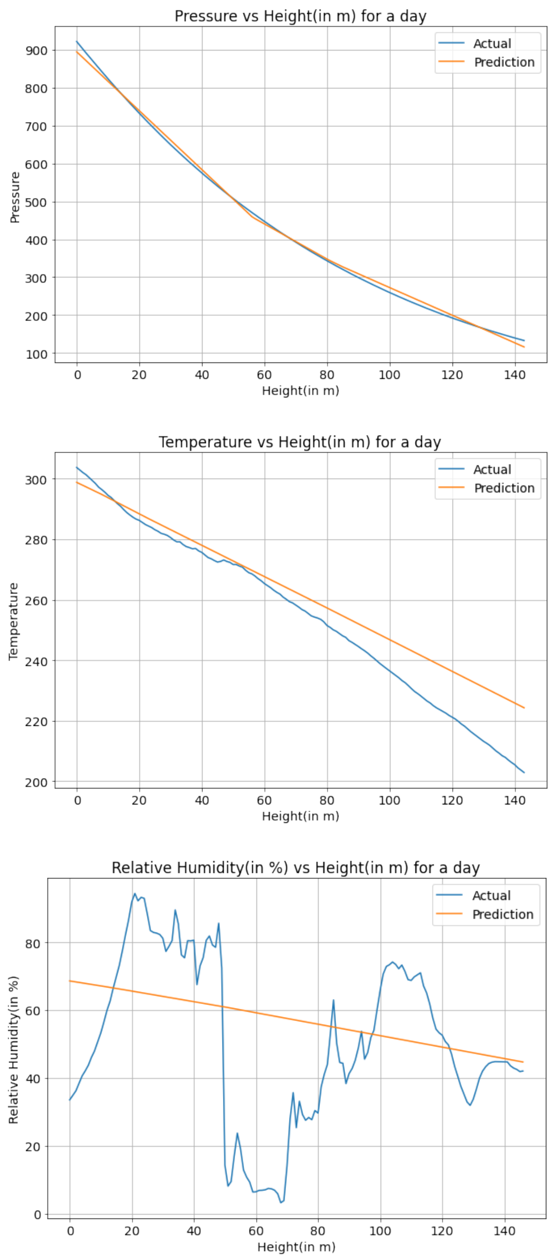



Figure 11: Wavelet Neural Network Graphs

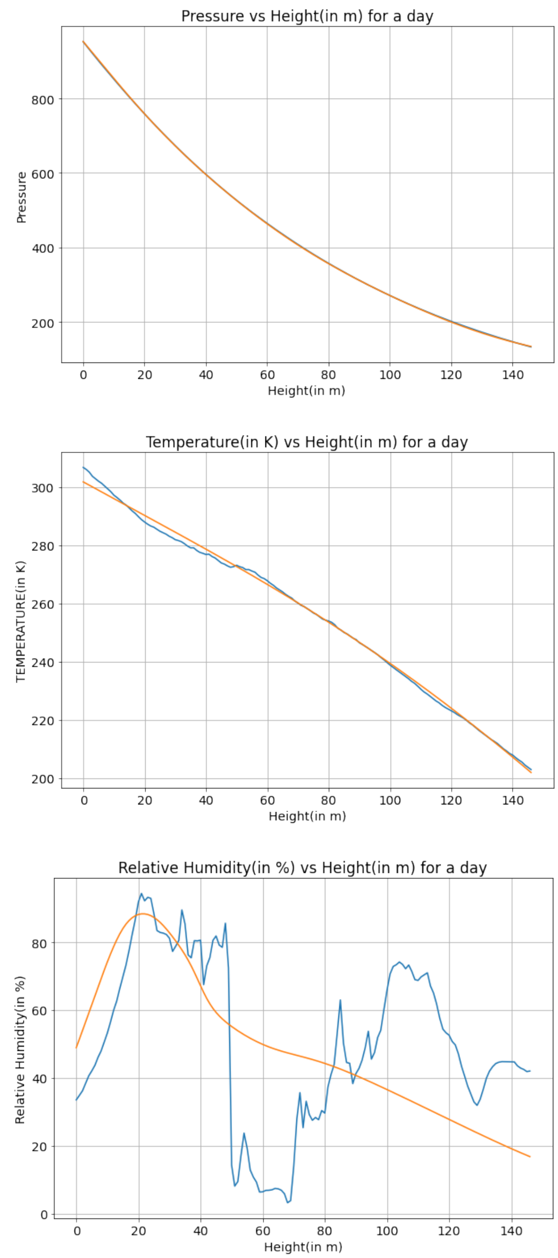



Figure 12: Modified Wavelet Neural Net Graphs

decomposition along with the introduction of Long Short Term Memory (LSTM) in the network helped the model adapt to the data better capturing the non-linearity in the mapping and addressing the underfitting problem.

## VII. Conclusion and Future Scope

In this study, a MWCNN was created for the recovery of pressure, temperature, and relative humidity profiles from ground-based radiometric measurements. In validation with radiosonde measurements, the results obtained by MWCNN showed a good retrieval capability with an percentage error of 0.37% for pressure, 0.91% for temperature, and 40.53% for relative humidity. Additionally, using the same training and test data, the effectiveness of the various retrieval techniques was compared with MWCNN wherein it outperformed all the other algorithms in all aspects. This research showed that retrieving vertical atmospheric profiles using MWCNN is a good solution.

There is potential for improving the model further especially for predictions of relative humidity by taking a deeper look into the data, their wavelet transforms and relation between them. The usage of different wavelets can also be explored to observe the effect it has in the learning procedure. Currently we are only working with a level-2 decomposition of the input, we can expand this to higher levels and can also study each sub-band to gain more insights. The architecture of the model also holds potential to be refined.

We also made a few attempts at trying the vector-to-vector mapping without the height parameter instead of the vector to a scalar(as sir had suggested). We used level-1 decomposition of the 22-channel data(11-Analysis Low Pass,11-Analysis High Pass) to predict the output which yielded a very high bias and hence a very poor accuracy. This needs further debugging and scrutinization to make it work and there is scope to try various combinations of wavelet decomposition to see if there is some correlation between different sub-bands of input and output.

## VIII. References


[1] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674-693, July 1989, doi: 10.1109/34.192463.

[2] Boureau, Y-Lan Ponce, J. Lecun, Yann. (2010). A Theoretical Analysis of Feature Pooling in Visual Recognition. ICML 2010 - Proceedings, 27th International Conference on Machine Learning. 111-118.

[3] C. -Y. Lee, P. Gallagher and Z. Tu, "Generalizing Pooling Functions in CNNs: Mixed, Gated, and Tree," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 863-875, 1 April 2018, doi: 10.1109/TPAMI.2017.2703082.

[4] T. Williams and R. Li, "Advanced Image Classification Using Wavelets and Convolutional Neural Networks," 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016, pp. 233-239, doi: 10.1109/ICMLA.2016.0046.

[5] Yu, D., Wang, H., Chen, P., Wei, Z. (2014). Mixed Pooling for Convolutional Neural Networks. In: Miao, D., Pedrycz, W., Ślzak, D., Peters, G., Hu, Q., Wang, R. (eds) Rough Sets and Knowledge Technology. RSKT 2014. Lecture Notes in Computer Science(), vol 8818. Springer.

[6] E. H. Shiguemori, H. F. De Campos Velho, J. D. S. Da Silva, and J. C. Carvalho, "Neural network based models in the inversion of temperature vertical profiles from radiation data," Inverse Problems Sci. Eng., vol. 14, no. 5, pp. 543–556, Jul. 2006.

[7] F. Solheim et al., "Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods," Radio Sci., vol. 33, no. 2, pp. 393–404, Mar. 1998.

[8] Y. Wang, Z. Wang, Q. Li, and Y. Zhu, "Research of the one-dimensional variational algorithm for retrieving temperature and humidity profiles from the ground-based microwave radiometer," Acta Meteorologica Sinica, vol. 72, no. 3, pp. 570–582, 2014.

[9] X. Yan, C. Liang, Y. Jiang, N. Luo, Z. Zang and Z. Li, "A Deep Learning Approach to Improve the Retrieval of Temperature and Humidity Profiles From a Ground-Based Microwave Radiometer," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 12, pp. 8427-8437, Dec. 2020, doi: 10.1109/TGRS.2020.2987896.

[10] M. P. Cadeddu, G. E. Peckham and C. Gaffard, "The vertical resolution of ground-based microwave radiometers analyzed through a multiresolution wavelet technique," in IEEE Transactions on Geoscience and Remote Sensing, vol. 40, no. 3, pp. 531-540, March 2002, doi: 10.1109/TGRS.2002.1000313.

[11] Williams, Travis Li, Robert. (2018). Wavelet Pooling for Convolutional Neural Networks.

[12] Krizhevsky, A. (2009), 'Learning Multiple Layers of Features from Tiny Images', , 32–33.

[13] Cireşan, Dan Meier, Ueli Masci, Jonathan Gambardella, Luca Maria Schmidhuber, Jürgen. (2011). High-Performance Neural Networks for Visual Object Classification. Computing Research Repository - CORR.