Atmos. Meas. Tech. Discuss., 7, C2186–C2196, 2014
www.atmos-meas-tech-discuss.net/7/C2186/2014/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Adaptive neuro fuzzy inference system for profiling of the atmosphere" *by* K. Ramesh et al.

K. Ramesh et al.

amit@narl.gov.in

Received and published: 15 August 2014

Reply to Interactive comment on "Adaptive neuro fuzzy inference system for profiling of the atmosphere" by K. Ramesh et al.

Anonymous Referee #1 Received and published: 15 May 2014

Note: As suggested by honorable anonymous referee 3, we have separated non rainy days microwave radiometer brightness temperatures and trained ANFIS again to see the impact of clouds on training of ANFIS system. In revised manuscript we have referred as ANFIS(NRD) i.e. ANFIS trained using microwave brightness temperature inputs only on non-rainy days observed during June-September 2011. The ANFIS presented in earlier manuscript is referred in revised manuscript as ANFIS(RD+NRD)

i.e. ANFIS trained using microwave brightness temperature inputs on rainy and non-rainy days observed during June-September 2011. The all figures are plotted again and shown in revised manuscript. We have modified the revised manuscript section according by taking views and suggestions of both honorable referees.

1. Comment: GENERAL COMMENTS The paper outlines an enhancement of adaptive neural network techniques for the retrieval of temperature and humidity profiles from ground-based microwave sounders by means of a fuzzy inference system. The paper is mostly well written and the topic fits well into the scope of AMT. Answer: Authors are thankful to honorable anonymous referee for this comment. The revised manuscript is modified as per suggestion by honorable referee 1.

2. Comment: However, the description of the developed ANFIS method is lacking and it would be very difficult to replicate the system based on the given details. Answer: The section 3 of revised manuscript has been rewritten with more clarity with suitable references. We have redrawn Fig. 2 to depict the structure of the ANFIS used in this manuscript with more clarity.

3. Comment The analysis of the retrieval quality is also lacking compared to literature. For example it is difficult to judge the bias and distribution shape of the values at different altitude levels. Answer: The retrieval quality compared with the literature and additional figures added to the manuscript as suggested by honorable referee 1 to judge the bias at different altitude levels along with the scatter plots as suggested in the comment no. 6 below. The session 1 of revised manuscript (Introduction) and Figures are submitted along with this reply.

4. Comment: page 2717: The whole introduction is devoid of references to literature. This needs to be corrected. Answer: We have rewritten the introduction section of the revised manuscript and incorporated literature survey related to retrieval of atmospheric parameters using microwave radiometer profiler. The modified Introduction and cited references are included in the document submitted as supplement to this reply.

5. Comment: page 2718, lines 3-6: The choice of ANN/ANFIS method over others, such as linear regression, optimal estimation or even 1D-VAR is not motivated. For example, the cited paper of Cimini et al. (2006) shows that a rather straightforward regression scheme (M-REG) following Rodgers (2000) produces less biased values over large altitude ranges. Answer: As suggested by honorable referee 1, the paragraph related to advantages of ANN/ANFIS over linear regression method is included in the Section 1 (Introduction) of revised manuscript. This session is enclosed in the supplementary document.

6. Comment: page 2723, Section 3.3: The choice of fitness quantities is peculiar. I am not sure why not simply the cited paper of Cimini et al. (2006) is followed and the bias and standard deviation are plotted (either in absolute or relative form). Especially the bias is relevant for judging the retrieval quality as this determines the gain achievable by averaging. The definition of rather well known diagnostics takes too much room (certainly compared to the terse ANFIS description, which the typical reader of this journal will be much less familiar with). It would be better to simply follow the approach in Cimini et al. and extend the following discussion. Answer: As suggested by honorable referee 1, we have followed the analysis of Cimini et al. (2006) and added two additional figures in the revised manuscript. These figures depict the scatter plot of the difference between ANFIS/ANN retrieved profiles with radiosonde (GPSsonde) profiles at 10 different altitudes Vs the actual predictions by ANFIS and ANN method. As suggested by honorable referee 3 we have removed original Fig 3 from revised manuscript and these new figures are listed as Fig. 3 a and 3 b. The quantitative values of bias and correlations at each retrieved altitude are mentioned in the Figure and discussed in the literature. The figures are enclosed along with answers to referee comments.

7. Comment Lastly, the achieved vertical resolution is not discussed at all. Answer: In this work, we desire to establish the ANFIS technique to retrieve the temperatures and relative humidity at 10 vertical locations at resolution of 1 Km. As aim of this work was to establish the ANFIS techniques for microwave radiometric profiler, we

have not addressed the issue of increasing vertical resolution based on the dataset. Also, vertical resolution of radiosonde data available for this study during observational campaign is of 250 m resolution. The availability of high resolution dataset for this study will enable the study of retrieval accuracy with increase in vertical resolution. The discussion related to this topic, as suggested by honorable referee 1, is included in the revised manuscript.

8. Comment Further, I am interested how many degrees of freedom the resulting profile contains, if this approach from optimal estimation is also applicable here (if it isn't this would seem to be a disadvantage of this method). Answer: The degrees of freedom of the total ANFIS system is 9 as we are using 10 parameters to retrieve 1 output i.e. brightness temperatures for 10 different frequencies observed by microwave radiometers to produce RH/Temperature for each level. Therefore degrees of freedom of ANFIS system are 9.

9. Comment: page 2725, Section 4.1: It is unclear why the radiosonde data is available more often at certain altitudes that differ between temperature and humidity. While radiosonde measurements are often compressed to only indicate significant changes, it seems very important to use high resolution data with at least one measurement per target altitude and radiosonde ascent. See for example Schneider et al. "An empirical study on the importance of a speed-dependent Voigt line shape model for tropospheric water vapor profile remote sensing", jqsrt, 2011, (doi:10.1016/j.jqsrt.2010.09.008) for an example on how to treat and compare against high-resolution radiosonde data. If the unavailability of data at certain altitudes is a necessary technical limitation of the available radiosondes or available dataset, it must be discussed. As shown in Schneider et al., the variability of humidity is also much higher than can be reproduced from ground-based measurements, which they compensate for by folding the radiosonde measurements with the averaging kernel of their retrieval system. While the ANFIS almost certainly cannot be easily quantified with a similar matrix, the retrieval has almost certainly also smoothing qualities due to the nature of the instrument sensitivity.

It should be discussed how this affects the training and validation. Answer: Authors are thankful to honorable referee 1, for very valuable suggestion. As discussed earlier we have retrieved the temperatures and relative humidity at 10 different heights with resolution of 1 Km. The aim of this work was to establish the ANFIS technique to retrieve the temperature and humidity of using ANFIS method and compare them with the available observations. Also, radiosonde dataset available for this study is of 250 m resolution. We think that the availability of high resolution dataset for the study will help to increase the vertical resolution of retrieval. Therefore, we think a separate study needs to be carried out to increase the vertical resolution of retrieval. However, as suggested by honorable referee 1, we have discussed about it in section 3 (Method).

10. Comment: page 2725, Section 4.2: The “test” profiles are selected to be interspersed in the training data. This choice should be discussed as well as how the system is envisioned to be used in practice. Shall the training data of one year be used to retrieve profiles of succeeding years, or shall the ANFIS system be constantly trained with 12:00 data to reliably retrieve temperature and humidity in the times in between regular radiosonde ascents? Answer: The datasets used for training and testing the ANFIS retrievals are described in detailed in revised manuscript. The period of the observations used for training and validation ANFIS model is from June to September 2011. The ANFIS model validated for 30 days viz. June 24-30, July 21-31, August 26-31 and September 26-30, 2011. The brightness temperature data observed from MWR for all the days of monsoon months (June-September) of the year 2011 at 12:00 UTC each day, excluding dates selected for validation are used as an input to ANFIS system in training phase. As regular profiles of radiosonde are available at each 12 UTC, the ANFIS system trained for 12 UTC observations. ANFIS system would have been more robust if it would have been trained using many radiosonde observations at regular interval each day. Unfortunately the obtaining periodic profiles of radiosonde at regular interval for long duration (monsoon months) to train ANFIS system are not economically feasible.

11. Comment: page 2726, line 13: It remains to be shown that spatial inhomogeniety of water vapour is the underlying cause for disagreement between retrieved and measured humidity profiles. All kind of systematic effects might be responsible, in particular rain. For example, the situation around 2011-07-27 has a bad correlation for several consecutive days, whereby the ANFIS is notably better. Is is unlikely that the ANFIS is capable of compensation for horizontal humidity gradients, so horizontal inhomogeniety cannot be the sole explanation. Can horizontal inhomogeniety be shown to be the real cause for an exemplary days using meteorological (model?) data? For example by showing that the days with bad correlation had especially high wind speeds? Answer: In this statement was written to indicate that even though radiosonde drifts far away from the radiometer station the change in temperature is relatively smaller compare to that in humidity field. As indicated by honorable referee 1, because this statement creates confusion in readers mind it is removed from the manuscript.

12. Comment: page 2727, Section 4.3: The behavior of SMAPE in Fig 5d seems to indicate a much larger bias for the ANFIS method, which does not seem to be supported by the Fig 5a. As the temperature value in a certain altitude remain in a rather small interval, a SMAPE of 3 percent should indicate a mean difference of about 2K, which cannot be discerned in Fig. 5a. A plot of the mean difference (bias) would certainly help diagnosing this. Answer: The scatter plot and trend lines for ANN and ANFIS models are shown in Fig 3 a, b. The values of bias calculated and mentioned in the Fig. 3 a, b. As seen from this analysis is that at lower altitudes the values of bias of ANFIS systems is much smaller compared to ANN however over middle altitudes ANN perform better. The discussion related to changes in SMAPE and bias is included in revised manuscript.

13. Comment: page 2728, Section 4.4: The mean profile of ANFIS shows much less bias than the ANN one. As the RMSE of both methods is comparable, this would suggest that the variance of ANFIS would be much larger. Lastly the SMAPE plot seemingly contradicts the mean plot, as the ANN model has an obvious and rather sig-

nificant dry bias at higher altitudes. This needs to be corrected or sufficiently explained. Answer: The explanation regarding dry bias at higher altitudes is added in the revised manuscript.

14. Comment: MINOR/TECHNICAL CORRECTIONS page 2718, line 21: “The ANN used in this MWR is useful to train vertical profiles...” This sentence doesn’t make sense to me. Isn’t the ANN trained by vertical profiles? Answer: The sentence is modified as below

“The ANN used in this MWR is trained using vertical profiles of the temperature and relative humidity observed by GPSsonde observation at the site. The input for this training with input as microwave radiances and the ANFISs are trained using the radiosonde observations as output.”

15. Comment: page 2719, line 10: FIS may incorporate human knowledge, which is one of its main advantages. In what way was human knowledge employed in building rules for the ANFIS? Answer: Authors agree that FIS incorporate human knowledge. ANFIS uses mathematical representation FIS is proposed by Takagi and Sugeno, 1983 to incorporate the knowledge from the dataset. In this work the Sugeno type subtractive clustering (Chiu 1997, Yager and Filev 1994) used for incorporating knowledge from input (microwave brightness temperatures) and output (radiosonde profiles). This method helps in reducing number of rules automatically determines number of clusters.

Takagi, T. and Sugeno, M.: Derivation of fuzzy control rules from human operator’s control action, in proc. IFAC Symp. Fuzzy inform., Knowledge Representation and Decision Analysis, pp. 55-60, July 1983. J.-S. R. Jang, C. T. Sun, and E. Mizutani, Neuro–Fuzzy and Soft Computing. Upper Saddle River, NJ: Prentice-Hall, 1997, ser. MATLAB Curriculum Series. Stephen L. Chiu, Fuzzy Information Engineering: A Guided Tour of Applications, ed. D. Dubois, H. Prade, and R. Yager, John Wiley & Sons, 1997. Yager, R. and D. Filev, "Generation of Fuzzy Rules by Mountain Clustering," Journal of Intelligent & Fuzzy Systems, Vol. 2, No. 3, pp. 209-219, 1994.

16. Comment: page 2720, line 9 - 11: I am not sure how to interpret the given vertical resolution figures for the MWR. As I fathomed the MWR to be a passive radiometer, upwards pointing measurements have, per se, no vertical sampling. However, as different channels have different altitude-varying sensitivities to temperature, pressure, and trace gas concentrations, one might deduce vertical profiles of these entities with certain resolutions. Answer: As discussed earlier we have selected the altitudes from 1 to 10 Km with resolution of 1 Km to train different ANFIS models for each altitude for each parameter. Therefore this aspect not discussed in the earlier manuscript. The discussion related to altitude varying sensitivities to temperature, pressure and trace gases and increasing vertical resolution of retrieved profiles is discussed in the revised manuscript.

17. Comment: page 2720, line 14 - 15: “These channels were selected based on their sensitivity to the occurrence of thunderstorms over the study site”. It is not clear, what makes these channels sensitive to thunderstorms. It is possible to, for example, determine the sensitivity of the measured signal with respect to changes in temperature or humidity in certain altitudes and optimize the channel selection accordingly. Or even better to optimize the information content or degrees of freedom of the retrieved profiles. Answer: The microwave radiometer available at NARL operates at 31 frequencies at Ka and V bands. However the training of ANFISs using all microwave brightness temperatures at all frequencies was difficult job. Therefore, we have studied the time series of variation of brightness temperatures observed by microwave radiometer on the days of thunderstorms. We found 10 channels selected in this study shows increase in magnitude of 3 - 4 hours prior to occurrence of thunderstorms. This aspect discussed in revised manuscript in detail and we have included plots of variation of brightness temperatures on thunderstorm day of May 28, 2013 for clear description of selection of input to ANFISs.

18. Comment: page 2720 line 18 - 20: Later on, it is related to missing training data at certain altitudes. You should describe here at what vertical sampling the radiosonde

data is given and how that sampling is converted to the sampling (regular 1 km grid) required by the ANFIS model Answer: As discussed earlier the description is added to revised manuscript.

19. Comment: page 2721, line 19: “Most of the rule-based prediction models need a few rules to predict”. This sentence seems rather pointless. Answer: Sentence removed from the revised manuscript

20. Comment: page 2721, line 20 - 21: “Since the number of predictors (10) is large, it may produce many dispiriting ANFIS structurs”. The number may be large for ANFIS models, but 10 measurements is not much for many nadir or limb-pointing satellites. Assuming that the measurements are well characterized with respect to accuracy and precision, adding more measurements should usually improve the result. Further, I do not know what “dispiriting” means in this context. Answer: The sentence is corrected as below. As mentioned above the discussion on selection of microwave brightness channels is added in the revised manuscript. The word “dispiriting” removed from the manuscript.

21. Comment: page 2722, line 14 - 19: This description follows uncomfortably close to Jangs 1993 paper. What is confusing in this context, is that Jangs paper explicitely notes that his example has two input with two membership functions each. This paper has 10 inputs and does not seem to specify how many membership functions are specified for each of the 10 inputs. It is also not clear, if the a_i and b_i factors are defined by humans or optimized by the ANFIS. Answer: The detailed structure of ANFIS and subtractive clustering for present ANFIS models are included in revised manuscript.

22. Comment: page 2722, line 23 (Eq. (3)): Here, the membership function B_i was not defined and it is unclear why the index i runs from 1 to 2, except that it did so in Jangs paper. It would be more helpful here to use the numbers and notions of the concrete ANFIS model described by this paper. As all this may be fully or partly automated by the subtractive fuzzy clustering, at least the initial set of rules should be

described concisely. The Fig.2 seems to imply that for each input many (21?) membership functions exists, whereby one membership function of each input is multiplicatively combined with one of all other inputs. It might also be instructive to show or discuss the structure of the reduced models, assuming that the structure of each model is different and somehow interpretable for each altitude level. Answer: The detailed architecture of the model will be described in revised manuscript as suggested by referee.

23. Comment: page 2723, line 6 (Eq. (5)): The use of f_i here seems to conflict with the use of f_i for observed values in Eq. (7). Also, comparing with Jang (1993) Eq. (22), the w_i on the r.h.s. seems to be lacking a bar. Answer: Corrected in revised manuscript.

24. Comment: page 2723, line 10: The meaning of this sentence is unclear. “more” than what? And why are a small number of rules needed? Answer: The sentence corrected as below. In this work, the numbers of predictor selected for development of ANFIS are 10. We have used Sugeno type subtractive fuzzy clustering (Chiu, 1994) to reduce the number of predictors to decrease the training rules in FIS to make ANFIS more robust. 25. Comment: page 2724, line 8: The definition of e_i should be e_i = f_i - y_i without the absolute value, else the SMAPE could not become negative as depicted later on. Answer: As suggested by honorable referee 1, we have corrected in revised manuscript by specifically mentioning this below SMAPE formula.

26. Comment: page 2726, line 2: The variable “r” should be introduced in some manner, even though it is frequently used for correlations. Answer: The more detailed description related to Pearson product movement correlation coefficient “r” is added in the section 3.3 Fitness of ANFIS models.

27. Comment: page 2727, line 2: I assume this relates to the profiles retrieved only from the test data. Answer: Yes these diagrams are related to retrieved profiles of test data. The clear description of data selected for the testing data and training data is added in the section 2 (data).

28. Comment: page 2727, line 10: “MSE” -> “MAE” ? Answer: Abbreviation corrected to MAE

29. Comment: page 2727, line 24 - 25: A reduction of the RMSE does not necessarily imply a reduction of bias, as the RMSE is the sum of the mean error (bias) squared and the error variance. The SMAPE plot is the closest diagnostic depicting something like a bias and ANFIS seems to be quite bad compared to ANN. Answer: We have calculated the bias for both parameters (temperature and RH) for all levels (as suggested by honorable referee 1 in comment no. 6) and shown in new figures 3 (a and b). It is seen from the figures that there is significant reduction in bias for ANFIS (RD+NRD) retrieval algorithum compared to ANN and ANFIS(NRD) retrieval algorithm. However, it is seen from the analysis that ANN has more systematic bias compared to both of other algorithms. The paragraph on this aspect is added to section 4: results and discussions.

30. Comment: page 2728, line 5: Fig. 6b should show the retrieved test data, not the trainings data, which is depicted in Fig. 3 Answer: As suggested by honorable referee 3, we have removed Fig.3 from the manuscript. Therefore we have modified the sentence as below. The computed RMSE of retrieved humidity averaged over training dataset was less than 0.01 % throughout the profile, however it is seen from Fig. 6b that the values of RMSE of the testing dataset for both (ANFIS(RD+NRD) and ANFIS(NRD)) models varies significantly with respect to height from 1-20%.

Please also note the supplement to this comment:
http://www.atmos-meas-tech-discuss.net/7/C2186/2014/amtd-7-C2186-2014-supplement.pdf

---

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 2715, 2014.