Tellus SERIES B CHEMICAL AND PHYSICAL METEOROLOGY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tracer–tracer relation in the Arctic stratosphere using fuzzy logic

By K. SATHEESAN*, *National Centre for Antarctic and Ocean Research, Vasco-da-Gama, Goa, India*

(Manuscript received 5 January 2012; in final form 20 August 2012)

ABSTRACT

Potential of finding a robust tracer–tracer relation using an approach based on fuzzy logic in the Arctic stratosphere using satellite observations of O_3 and N_2O has been demonstrated. Using this tracer–tracer relation, O_3 is predicted in terms of N_2O, latitude, longitude, altitude and day of the year. The model for O_3 based on this approach has been validated with independent observations and has been found to be superior to those based on regression. This method has the potential to generate long time series of chemical species data from different satellite sensors in a self-consistent manner. These data have use in model evaluations and data assimilation.

Keywords: stratosphere, arctic, trace gases, tracer relation, fuzzy logic

1. Introduction

The chemical and radiative balance of the stratosphere is dependent on many long-lived chemical species. The spatial distribution of long-lived chemical species in the stratosphere is dependent on both chemistry and transport. Chemistry–climate models need to be validated against observations for checking how well they represent the processes in order to have confidence in the model predictions of climate change and different scenarios. Unfortunately, a sufficient number of observations of many chemical species are not available for validation of the models. Satellite observations of chemical species are useful for validating these models even though the observations are not homogeneous in space and time. Combining measurements from multiple satellites will improve the spatial and temporal coverage. Randel and Wu (2007) generated a data set of O_3 for 1979–2005 by combining different satellites. Combining measurements from different satellites requires that measurements of same species be available from different satellites.

Sufficiently long-lived species in the stratosphere exhibit compact correlations (Plumb and Ko, 1992), which eliminate day-to-day variations and give a means to provide a robust constraint on models. The relationships between O_3 and N_2O and between O_3 and CH_4 are usually used in model comparisons and as diagnostics in the transport and mixing in the stratosphere. Even though O_3 is not a long-lived tracer in the whole of the stratosphere, the photochemical lifetime of O_3 can be up to 1 yr within the polar vortex in the stratosphere. The O_3–N_2O correlation is used in many studies of transport and mixing of the stratosphere (Waugh et al., 1997; Bregman et al., 2000); an anomaly in the linear correlation curve between these two is interpreted as mixing between polar vortex and the mid-latitude environment (Waugh et al., 1997). The correlations between different trace constituents are also used to find stratospheric flux of O_3 into the troposphere (Murphy and Fahey, 1994), seasonal evolution of O_3 and N_2O (Khosrawi et al., 2008) and so on. Similar strong correlations exist for many long-lived tracers due to their transport by the general circulation of the atmosphere even if they are not related chemically.

The tight relationships between different constituents have led to many analyses where measurements of one tracer are used to infer the abundance of another tracer. In other words, the abundance of one tracer can be modelled in terms of another tracer by making use of their spatial and temporal correlations. The concentration of any one of the constituents can be predicted from the knowledge of the concentration of a single reference species (Plumb, 2007). Such models of chemical species can be employed to generate long time series data of chemical species by combining different satellites. For example, if we have a model of N_2O determined by O_3, data sets of N_2O can be generated using observations of O_3 from satellites

*Corresponding author.
email: satheesan.k@gmail.com

Citation: Tellus B 2012, **64**, 17164, http://dx.doi.org/10.3402/tellusb.v64i0.17164

such as High Resolution Dynamics Limb Sounder (HIRDLS) which currently does not give N_2O. The relationships between different trace constituents vary spatially and temporally because of the latitudinal and seasonal mixing; therefore, linear relationships will not hold under all regional and seasonal conditions. Linear correlation plots will only help in identifying the areas of mixing and transport. Modelling of such relations, which are usually governed by complex non-linear processes, can be achieved empirically by making use of the available data. Powerful optimisation tools are usually required for modelling such complex processes in order to find a relationship between different variables (e.g. Satheesan and Kirkwood, 2010). Non-linear techniques such as genetic algorithm, artificial neural network (ANN) and so on have been found to be successful in many geophysical problems (e.g. Lary et al., 2004; Basu et al., 2005a, b). These methods derive mathematical models which are not based on physical processes, but by analysing a system and finding connections between the variables without explicit knowledge of the physical behaviour of the system. Lary et al. (2004) used ANN to describe tracer correlations between CH_4 and N_2O using data obtained from a chemical model. In this study, a non-linear method based on a combination of fuzzy logic and subtractive clustering techniques is used to find a relation between O_3 and N_2O. Fuzzy-logic-based methods are one of the major techniques successful in non-linear system identification used in many areas such as communication, control systems, signal processing, chemical process control, biological processes, atmospheric parameter retrievals, turbulence classification etc (Sugeno, 1985; Center and Verma, 1998; Kottayil et al., 2010; Satheesan and Kirkwood, 2010). In this work, the potential of fuzzy logic combined with subtractive clustering for finding a robust relationship between O_3 and N_2O in the arctic stratosphere using satellite data is demonstrated. After a relation is thus established, we can use it for predicting the abundance of one of the tracers in terms of the other. I use this method here, to predict O_3 from the observations of N_2O.

2. Data

Chemical species data of O_3 and N_2O measured by the Improved Limb Atmospheric Spectrometer-II (ILAS-II) operated by the Japan Aerospace Exploration Agency (JAXA) has been used. Measurements of vertical profiles of O_3, N_2O and many other trace gases are available from ILAS-II with a vertical resolution of 1 km (Nakajima et al., 2006). The measurements made in the latitudes between 54°N to 71°N and in the altitude regions 10–50 km during April–October 2003 were used. The N_2O values derived from ILAS-II Version 1.4 agree with the balloons and Odin Sub-Millimeter Radiometer (SMR) within 20% when the N_2O volume mixing ratio (VMR) is greater than 250 ppbv. When N_2O VMR is between 30–50 and 250 ppbv, the ILAS-II data is underestimated by 10–30% (Ejiri et al., 2006). The O_3 data derived from ILAS-II agrees with other data within 10% at altitudes between 11 and 40 km with a positive bias below 20 km and a negative bias above it. The negative bias increases above 40 km and reaches 30% at 61–65 km (Sugita et al., 2006). In the altitude range considered in this study, the data is suitable for quantitative analyses in the high latitude region. The ILAS-II data used here (Version 2) has been improved compared to Version 1.4 by improving the transmittance correction in the Northern Hemisphere and by using an improved tangent height registration (Tanaka et al., 2007).

3. Methods

3.1. *Fuzzy logic*

Fuzzy logic is suitable for application in problems dealing with imprecision caused by the absence of sharp criteria. Advancement of fuzzy logic began with the introduction of fuzzy set theory by Zadeh (1965). Fuzzy sets (Zadeh, 1965) and fuzzy operators are the basic building blocks of fuzzy logic. A fuzzy set is a set without a crisp and clearly defined boundary and is an extension of the classical sets. How an element belongs to a particular fuzzy set is represented by a membership function whose value lies between 0 and 1. Application of fuzzy methods for solving geophysical problems is becoming more and more popular (e.g. Kottayil et al., 2010; Satheesan and Kirkwood, 2010). The approach (applied to model turbulence) followed by Satheesan and Kirkwood (2010) is adapted here in exploring the applicability of fuzzy methods to the tracer–tracer relation in the arctic stratosphere. Our main objective is to find a relation between O_3 and N_2O at specific locations and time. To find this, N_2O, latitude (q), longitude (f), altitude (Z) and the day number (N, day of the year starting from January 1) are used as the inputs and O_3 as the output (Satheesan and Kirkwood, 2010) in the Takagi–Sugeno model (Takagi and Sugeno, 1985) where IF-THEN rule statements are used to formulate the conditional statements that comprise the model (Angelov and Filev, 2004). There are two parts to the rules describing a model: the antecedent part between the IF and THEN, where the inputs are defined, and the consequent part following THEN, where the output is specified. A detailed explanation of the procedure followed in applying fuzzy logic is given by Satheesan and Kirkwood (2010).

There are different ways of achieving fuzzy modelling. One of them is to find an input–output relation of data with a fuzzy relation. When using this method, a large number of fuzzy sets will be required to cover the entire range of variation of the variables. This will create a large number of fuzzy rules because the number of rules is proportional to the number of input variables, which will increase the size of the rule base constituting the model. If the dimensionality of the problem can be reduced, it is possible to avoid a rule base with large size. This can be achieved by classifying the data into a pre-determined number of groups or clusters, using some clustering algorithms. Each cluster will be represented by its centre, thus by reducing the rule base to a relatively small number of rules. Therefore, clustering of the data is performed to bring out natural groupings from a large data set and to obtain a concise representation of a system's behaviour.

3.2. *Data clustering*

Data clustering will help in classification of data into similar categories or patterns. Clustering aids not only in classification and pattern recognition of the available data but also for the reduction of complexity in modelling and identification. A variety of data clustering algorithms are available such as hard c-means (HCM) and fuzzy c-means (FCM). These are supervised clustering algorithms, and it is essential to specify the number of clusters required in these methods. Unsupervised clustering algorithms are used if the number of clusters is not known beforehand or it cannot be specified. Subtractive clustering is an unsupervised clustering algorithm, and it is based on the density of data points in the feature space (Chiu, 1994). The aim of clustering is to find regions in the feature space with a high density of data points. In the subtractive clustering, the point with highest number of neighbours is selected as the centre for a cluster. The data points contained within this selected cluster are removed (subtracted) in order to make sure of its absence in the next cluster. After its removal, the algorithm will search for a new point with the next highest number of neighbours. This process is stopped after all the data points are assigned to one of the clusters.

The subtractive clustering method (Chiu, 1994; Jang et al., 1997) is used here since it is an unsupervised clustering algorithm with automatic rules generation capability. In this algorithm, each data point is considered as a potential cluster centre. Therefore, the number of points to be evaluated is independent of the dimension of the problem and is simply equal to the number of data points. The modelling of O_3 is achieved through subtractive data clustering and fuzzy logic. In the first step, the entire dataset is clustered into different homogeneous groups using the subtractive clustering algorithm. Following this, the fuzzy rules are applied to each cluster centre individually. This has an advantage that the number of fuzzy rules can be reduced significantly (equivalent to number of clusters), and the redundancy in the rules can be avoided. A search radius has to be specified for the subtractive clustering, which will determine the number of clusters generated. The search radius is the Euclidean distance between the cluster centre and the data points. Smaller radius will give a very good model for the training set since more number of points are used for training. But, this will not guarantee to be the best model and may not give good results for the validation using independent data sets. This problem is due to over-fitting. So, there must be a trade-off between the search radius and the desired model accuracy. This radius is determined depending on the characteristics of the variability of the dataset under consideration and has to be chosen by trial and error.

4. Results

The training data set consisted of O_3 and N_2O in the stratosphere (10–50 km). There were 47 872 data points available for this altitude range. Figure 1 shows the O_3–N_2O scatter plot for the entire data in stratosphere, available from ILAS-II in the Northern Hemisphere. It is obvious from the figure that the O_3–N_2O relations are not compact when we consider the whole of the stratosphere. The scatter shows the shape of a boomerang. The positive and negative correlation in the figure is from the region above and below the ozone maximum. Figure 2 shows the relationship of O_3 with altitude, longitude, latitude and

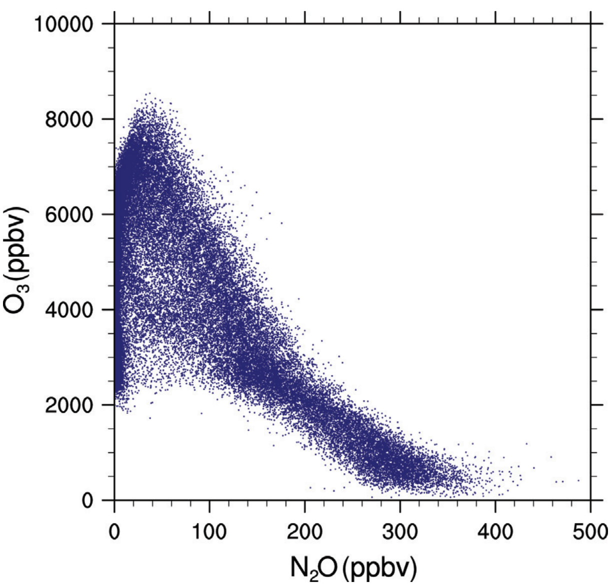

Fig. 1. Scatter plot of O_3 versus N_2O for the whole data. The scatter shows the shape of a boomerang. The positive and negative correlation in the figure is from the region above and below the ozone maximum.

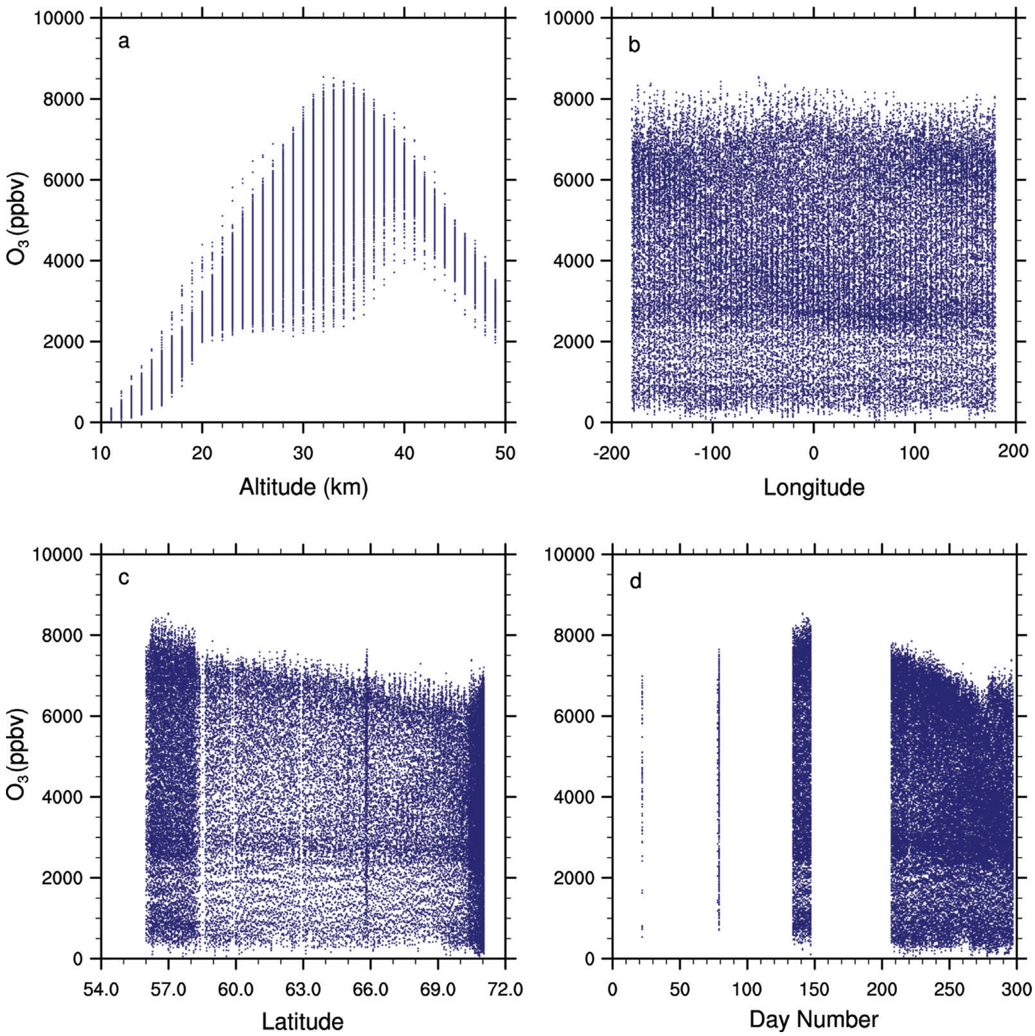

Fig. 2. Scatter plot of O_3 against a) altitude, b) longitude, c) latitude and d) day number.

the day number in the Northern Hemisphere stratosphere. The day number is calculated from January 1 of 2003, the year for which the data is available. It can be observed from Figure 2 that there is a strong relation for O_3 with altitude. The scatter plot has the shape of a boomerang with maximum value for O_3 in the mid-stratosphere. O_3 does not show much dependence on longitude but shows a weak relation with latitude and the day number. Therefore, any model of O_3 in the stratosphere should have altitude, latitude and day number as the predictors along with N_2O. Longitude is also added as one of the predictors since there is significant longitudinal variability for O_3 in the stratosphere in the high latitudes (Natarajan and Callis, 1997).

In the first step, the number of input variables was reduced using subtractive clustering so that the fuzzy rules are needed to be generated for less number of data. Out of the total 47872 data points available, only 23936 (50%) data points were used for training using fuzzy techniques. The remaining data points were used for validation. For clustering, different values of radii were chosen to find an optimum value so that the number of points is sufficient enough to give a good model and not too large to make it complex and cause over-fitting. The clusters for different choices of radii are shown in Figure 3. The figure is shown only for the cases of N_2O and altitude. The number of clusters obtained for the respective choices is also given in the figure. It is obvious that a small search radius will produce a larger number of clusters while a big search radius will generate a small number of clusters. A cluster centre represents the same physics as the points surrounding the cluster centre with less number of points. If we consider the clusters for the radius 0.3 (blue markers in Figure 3), all the values within the dynamic range of N_2O and O_3 are represented by the cluster centres.

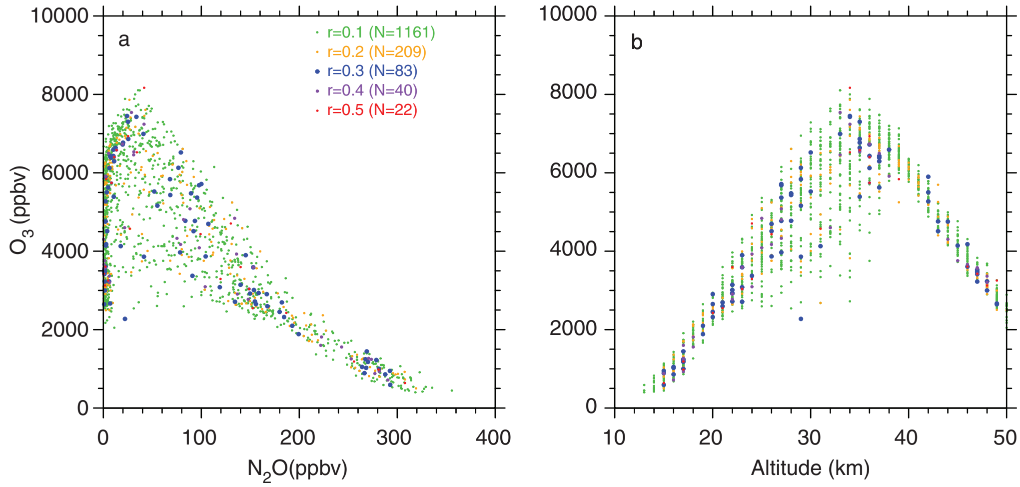

Fig. 3. The clusters for the choice of different search radii. The number of clusters obtained is shown within the bracket for each search radius. A search radius of 0.3 is used in the present study.

As per the procedure explained by Satheesan and Kirkwood (2010), the cluster centres, as input values, are graded between the values of 0 and 1. They are coded in the form of a fuzzy membership function of different shapes. The Gaussian curve membership function was chosen as they are most adequate in representing uncertainty in measurements (Kreinovich et al., 1992). IF–THEN rules by Takagi and Sugeno (1985) are used as the model to generate the rules. The output of each rule is a linear combination of input variables and a constant term. The final output will be the weighted average of each rule's output. Using a T-norm operator (multiplication or minimum), the premise part of each rule is combined to get the weight. A sensitivity test was conducted by varying the search radius for clustering to check the effect of the choice of radius on the model. The radius was varied from 0.1 to 0.5. When the radius was 0.5, the *R* obtained is 0.979 which increased to 0.986 when the radius was 0.4. Decreasing radius from 0.3 to 0.2 did not increase the *R* considerably even though the number of clusters more than doubled which will generate more fuzzy rules making the model more complex. Therefore, using the training set, fuzzy rules were generated using a search radius of 0.3 for clustering. The number of fuzzy rules was 83 which represent the data for the entire dynamic range. Figure 4a shows the scatter plot of the O_3 for the training data. The correlation coefficient (*R*) for training data is 0.99 with a root mean square error (RMSE) of 274. Figure 4b shows the same for the validation data. It has an *R* of 0.991 and RMSE of 239. The correlation is very high for both training

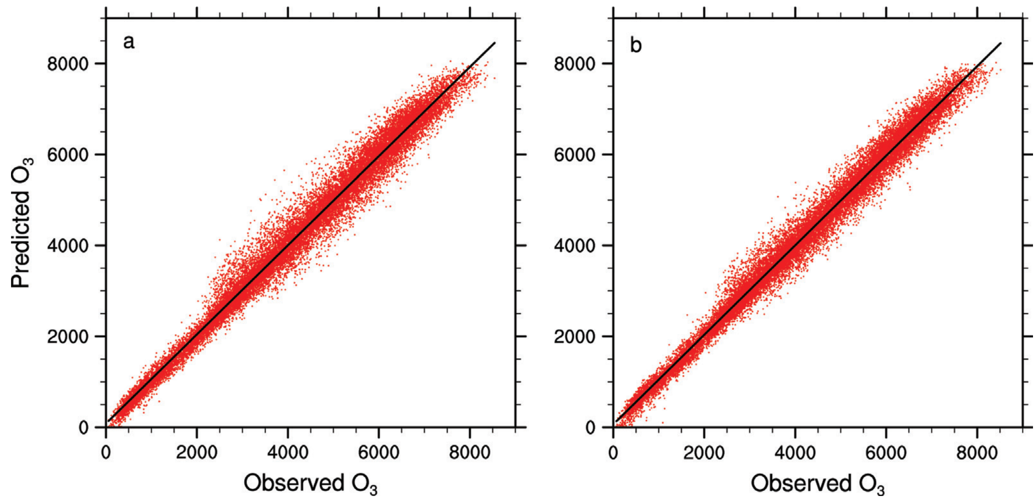

Fig. 4. O_3 (in ppbv) observed and predicted for a) training data and b) validation data set.

Table 1. Comparison of the performance the models based on fuzzy and regression techniques. The number of data points used was 23 936 for both training and validation

	Fuzzy				Linear			
	R^2	RMS	Norm	Slope	R^2	RMS	Norm	Slope
Training	0.990	274	42 395	1	0.843	1052	137 293	0.71
Validation	0.991	239	37 079	1	0.843	1052	136 971	0.71

and validation data sets. For the sake of comparison, the classical regression modelling approach is also used for modelling. In this method, if X_i are independent variables and Y is a dependent variable, a model relating Y to X_i and unknown parameters a_i in the form $Y = a_0 + \Sigma\, a_i\, X_i$ is found. The unknown values a_i are found by the method of least squares in which the distance between the measured and predicted values of the dependent variable Y is minimum.

A comparison between the fuzzy model and the regressive model has been summarized in Table 1. A surprising observation is that R has increased slightly and RMSE has decreased for the validation points. This result indicates that the model represented by the fuzzy membership functions represents O_3 in the region very well. Thus, the fuzzy method offers a very good solution to reconstruct the tracers from different satellites or from the ground-based observations. This method makes use of the fact that knowledge of the concentration of one tracer in the stratosphere can be used to predict the concentration of any of the other constituents (Plumb, 2007). However, the tracer correlations in the stratosphere are altitude dependent, and a family of correlations are usually required to describe the relations over the entire altitude range in stratosphere. Therefore usually, such models for prediction of the concentration of any one of the tracers in terms of another require the relations to be established at different spatio–temporal locations. In the case with fuzzy logic, there was no need to find the relation separately for different altitude regions.

The application of tracer–tracer relation for the modelling as described above is based on the assumption that the tracer relations are invariant over the time periods under consideration. There may be instances when the tracer relations changes at different timescales. This can happen when concentrations of some of the species vary due to changes in the chemical regime. On shorter timescales, this can happen due to mixing and transport of species through tropopause or between different regions like polar and mid-latitude. These changes are basically seasonal, and the time information given in the model as one of the predictors will take care of such situations. Over longer timescales, the relation between two tracers can change due to the difference in the residence time (age) of the tracers or due to a change in the concentration of one of the species as a result of climatic shift. However, there is a lack of information on how the tracer–tracer relation varies over long timescales. Nevertheless, a solution to check for the consistency of relation could be the use of multiple tracers, in combination or separately, to establish the relation. Multiple tracers can be used to obtain the information on the tracer transport times (Waugh et al., 2003). The Fuzzy model also needs to be re-trained with independent data sets from time to time. In any case, a data set generated using this method will have to be validated with independent observations obtained through balloons, aircrafts, rockets and so on before being used for long-term studies and trend analysis. These data sets will be useful in study of the variability of the stratosphere and in model validations.

5. Summary and outlook

Fuzzy logic in combination with subtractive clustering has been used for establishing a robust tracer–tracer relation between O_3 and N_2O in the Arctic stratosphere. Using this relation, O_3 is predicted in terms of N_2O and the spatio–temporal variables of latitude, longitude, altitude and the day number. The method has been tested using ILAS-II satellite data of these trace gases over the arctic stratosphere. This method finds a relation between O_3 with the predictants N_2O, latitude, longitude, altitude and the day number and has been validated using independent data with very high correlation. The entire stratosphere is represented by one relation which otherwise would have needed a family of correlations. The fuzzy logic modelling approach presented in this paper performs much better than the regression modelling approach.

This method can be applied to any combination of trace gases, which have sufficiently long chemical lifetime. This will help in generating long time series of chemical species in a self-consistent manner after taking sufficient care to account for any variabilities in the tracer relations over the time period under consideration. This study can be extended to different regions of the globe including Antarctica in order to generate self-consistent data sets of different chemical species. Such data sets have use in model validation, assimilation and climate studies.

6. Acknowledgements

The author thanks the Director, National Centre for Antarctic and Ocean Research (NCAOR), Goa, for supporting this study. The ILAS-II data were processed at the ILAS-II Data Handling Facility of the National Institute for Environmental Studies, Japan. The ILAS-II project was funded by the Ministry of the Environment, Japan. This is NCAOR contribution 29/2012.

References

Angelov, P. P. and Filev, D. P. 2004. An approach to online identification of Takagi-Sugeno fuzzy models. *IEEE Trans. Syst. Man Cybern. B Cybern.* **34**, 484–498.

Basu, S., Satheesan, K., Sarkar, A. and Kishtawal, C. M. 2005a. Ocean surface wind prediction in the north Indian Ocean from in situ and satellite observations using genetic algorithm. *Geophys. Res. Lett.* **322**, L24601. DOI: 10.1029/2005GL024827.

Basu, S., Sarkar, A., Satheesan, K. and Kishtawal, C. M. 2005b. Predicting wave heights in the north Indian Ocean using genetic algorithm. *Geophys. Res. Lett.* **321**, L17608. DOI: 10.1029/2005GL023697.

Bregman, A., Lelieveld, J., van den Broek, M. M. P., Siegmund, P. C., Fischer, H. and co-authors. 2000. N_2O and O_3 relationship in the lowermost stratosphere: A diagnostic for mixing processes as represented by a three-dimensional chemistry-transport model. *J. Geophys. Res.* **105**, 17279–17290. DOI: 10.1029/2000JD900035.

Center, B. and Verma, B. P. 1998. Fuzzy logic for biological and agricultural systems. *Artif. Intell. Rev.* **12**, 213–225.

Chiu, S. 1994. Fuzzy model identification based on cluster estimation. *J Intel. Fuzzy Syst.* **2**(3), 267–278.

Ejiri, M. K., Terao, Y., Sugita, T., Nakajima, H., Yokota, T. and co-authors. 2006. Validation of the Improved Limb Atmospheric Spectrometer-II (ILAS-II) Version 1.4 nitrous oxide and methane profiles. *J. Geophys. Res.* **111**, 22. DOI: 10.1029/2005JD006449.

Jang, J.-S. R., Sun, C.-T. and Mizutani, E. 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice-Hall, Upper Saddle River, NJ.

Khosrawi, F., Müller, R., Proffitt, M. H., Urban, J., Murtagh, D. and co-authors. 2008. Seasonal cycle of averages of nitrous oxide and ozone in the Northern and Southern Hemisphere polar, midlatitude, and tropical regions derived from ILAS/ILAS-II and Odin/SMR observations. *J. Geophys. Res.* **113**, D18305. DOI: 10.1029/2007JD009556.

Kottayil, A., Thapliyal, P. K., Shukla, M. V., Pal, P. K., Joshi, P. C. and co-authors. 2010. A new technique for temperature and humidity profile retrieval from infrared sounder observations using adaptive neuro-fuzzy inference system. *IEEE Trans. Geosci. Remote Sens.* **48**, 1650–1659. DOI: 10.1109/TGRS.2009.2037314.

Kreinovich, V., Quintana, C. and Reznik, L. 1992. Gaussian membership functions are most adequate in representing uncertainty in measurements. In: *Proceedings of the North American Fuzzy Information Processing Society Conference, Puerto Vallarta*, pp. 618–624. Online at: http://www.cs.utep.edu/vladik/1992/tr92-8.ps.gz

Lary, D. J., Müller, M. D. and Mussa, H. Y. 2004. Using neural networks to describe tracer correlations. *Atmos. Chem. Phys.* **4**, 143–146.

Murphy, D. M. and Fahey, D. W. 1994. An estimate of the flux of stratospheric reactive nitrogen and ozone into the troposphere. *J. Geophys. Res.* **99**, 5325–5332. DOI: 10.1029/93JD03558.

Nakajima, H., Sugita, T., Yokota, T., Ishigaki, T., Mogi, Y. and co-authors. 2006. Characteristics and performance of the Improved Limb Atmospheric Spectrometer-II (ILAS-II) on board the ADEOS-II satellite. *J. Geophys. Res.* **111**, D11S01. DOI: 10.1029/2005JD006334.

Natarajan, M. and Callis, L. B. 1997. Ozone variability in the high latitude summer stratosphere. *Geophys. Res. Lett.* **24**, 1191–1194. DOI: 10.1029/97GL01155.

Plumb, R. A. 2007. Tracer interrelationships in the stratosphere. *Rev. Geophys.* **45**, 4005. DOI: 10.1029/2005RG000179.

Plumb, R. A. and Ko, M. K. W. 1992. Interrelationships between mixing ratios of long-lived stratospheric constituents. *J. Geophys. Res.* **971**, 10145–10156.

Randel, W. J. and Wu, F. 2007. A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data. *J. Geophys. Res.* **112**, 6313. DOI: 10.1029/2006JD007339.

Satheesan, K. and Kirkwood, S. 2010. Turbulence for different background conditions using fuzzy logic and clustering. *Ann. Geophys.* **28**, 1475–1481. DOI: 10.5194/angeo-28-1475-2010.

Sugeno, M. 1985. *Industrial Applications of Fuzzy Control*. Elsevier, New York, NY pp.

Sugita, T., Nakajima, H., Yokota, T., Kanzawa, H., Gernandt, H. and co-authors. 2006. Ozone profiles in the high-latitude stratosphere and lower mesosphere measured by the Improved Limb Atmospheric Spectrometer (ILAS)-II: Comparison with other satellite sensors and ozonesondes. *J. Geophys. Res.-Atmos.* **111**, 11. DOI: 10.1029/2005JD006439.

Takagi, T. and Sugeno, M. 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **15**, 116–132.

Tanaka, T., Nakajima, H., Sugita, T., Ejiri, M. K., Irie, H. and co-authors. 2007. Tangent height registration method for the version 1.4 data retrieval algorithm of the solar occultation sensor ILAS-II. *Appl. Optic.* **46**, 7196–7201.

Waugh, D. W., Hall, T. M. and Haine, T. W. N. 2003. Relationships among tracer ages. *J. Geophys. Res.- Oceans.* **108**, 3138. DOI: 10.1029/2002JC001325.

Waugh, D. W., Plumb, R. A., Elkins, J. W., Fahey, D. W., Boering, P. A. and co-authors. 1997. Mixing of polar vortex air into middle latitudes as revealed by tracer-tracer scatterplots. *J. Geophys. Res.* **1021**, 13119–13134. DOI: 10.1029/96JD03715.

Zadeh, L. 1965. Fuzzy sets. *Inform Control.* **8**, 338–353. DOI: 10.1016/S0019-9958(65)90241-X.