

Using neural networks to describe tracer correlations

D. J. Lary^{1,2,3}, M. D. Müller^{1,4}, and H. Y. Mussa³

¹Global Modelling and Assimilation Office, NASA Goddard Space Flight Center, USA

²GEST at the University of Maryland Baltimore County, MD, USA

³Unilever Cambridge Centre, Department of Chemistry, University of Cambridge, UK

⁴National Research Council, Washington DC, USA

Received: 8 October 2003 – Published in Atmos. Chem. Phys. Discuss.: 13 November 2003

Revised: 13 January 2004 – Accepted: 25 January 2004 – Published: 31 January 2004

Abstract. Neural networks are ideally suited to describe the spatial and temporal dependence of tracer-tracer correlations. The neural network performs well even in regions where the correlations are less compact and normally a family of correlation curves would be required. For example, the CH₄-N₂O correlation can be well described using a neural network trained with the latitude, pressure, time of year, and CH₄ volume mixing ratio (v.m.r.). In this study a neural network using Quickprop learning and one hidden layer with eight nodes was able to reproduce the CH₄-N₂O correlation with a correlation coefficient between simulated and training values of 0.9995. Such an accurate representation of tracer-tracer correlations allows more use to be made of long-term datasets to constrain chemical models. Such as the dataset from the Halogen Occultation Experiment (HALOE) which has continuously observed CH₄ (but not N₂O) from 1991 till the present. The neural network Fortran code used is available for download.

1 Introduction

The spatial distributions of atmospheric trace constituents are in general dependent on both chemistry and transport. Compact correlations between long-lived species are well-observed features in the middle atmosphere, as for example described by Fahey et al. (1989); Plumb and Ko (1992); Loewenstein et al. (1993); Elkins et al. (1996); Keim et al. (1997); Michelson et al. (1998); Rinsland et al. (1999); Strahan (1999); Fischer et al. (2000); Muscari et al. (2003). The correlations exist for all long-lived tracers – not just those which are chemically related – due to their transport by the general circulation of the atmosphere. 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. These correlations can also be used as a diagnostic of mixing (Schoeberl et al., 1997; Morgenstern et al., 2002) and to distinguish between air-parcels of different origins (Waugh and Funatsu, 2003).

Of special interest are the so-called “long-lived tracers”: constituents such as nitrous oxide (N₂O), methane (CH₄), and the chlorofluorocarbons (CFCs) that have long lifetimes (many years) in the troposphere and lower stratosphere, but are destroyed rapidly in the middle and upper stratosphere.

The correlations are spatially and temporally dependent. For example, there is a “compact-relation” regime in the lower part of the stratosphere and an “altitude-dependent” regime above this. In the compact-relation region, the abundance of one tracer is uniquely determined by the value of the other tracer, without regard to other variables such as latitude or altitude. In the altitude-dependent regime, the correlation generally shows significant variation with altitude (Minschwaner et al., 1996) (Fig. 1d).

The description of such spatially and temporally dependent correlations are usually achieved by a family of correlations. However, a single neural network is a natural and effective alternative.

1.1 Reconstructing N₂O-CH₄ correlations

The motivation for this study was preparation for a long term chemical assimilation of Upper Atmosphere Research Satellite (UARS) (Reber et al., 1993) data starting in 1991 and coming up to the present. For this period we have continuous version 19 data from the Halogen Occultation Experiment (HALOE) (Russell et al., 1993) but not observations of N₂O as both ISAMS and CLAES failed. In addition we would like to constrain the total amount of reactive nitrogen, chlorine, and bromine in a self-consistent way (i.e. the correlations between the long-lived tracers is preserved). Tracer correlations provide a means to do this by using HALOE CH₄ observations.

Correspondence to: D. J. Lary
(David.Lary@umbc.edu)

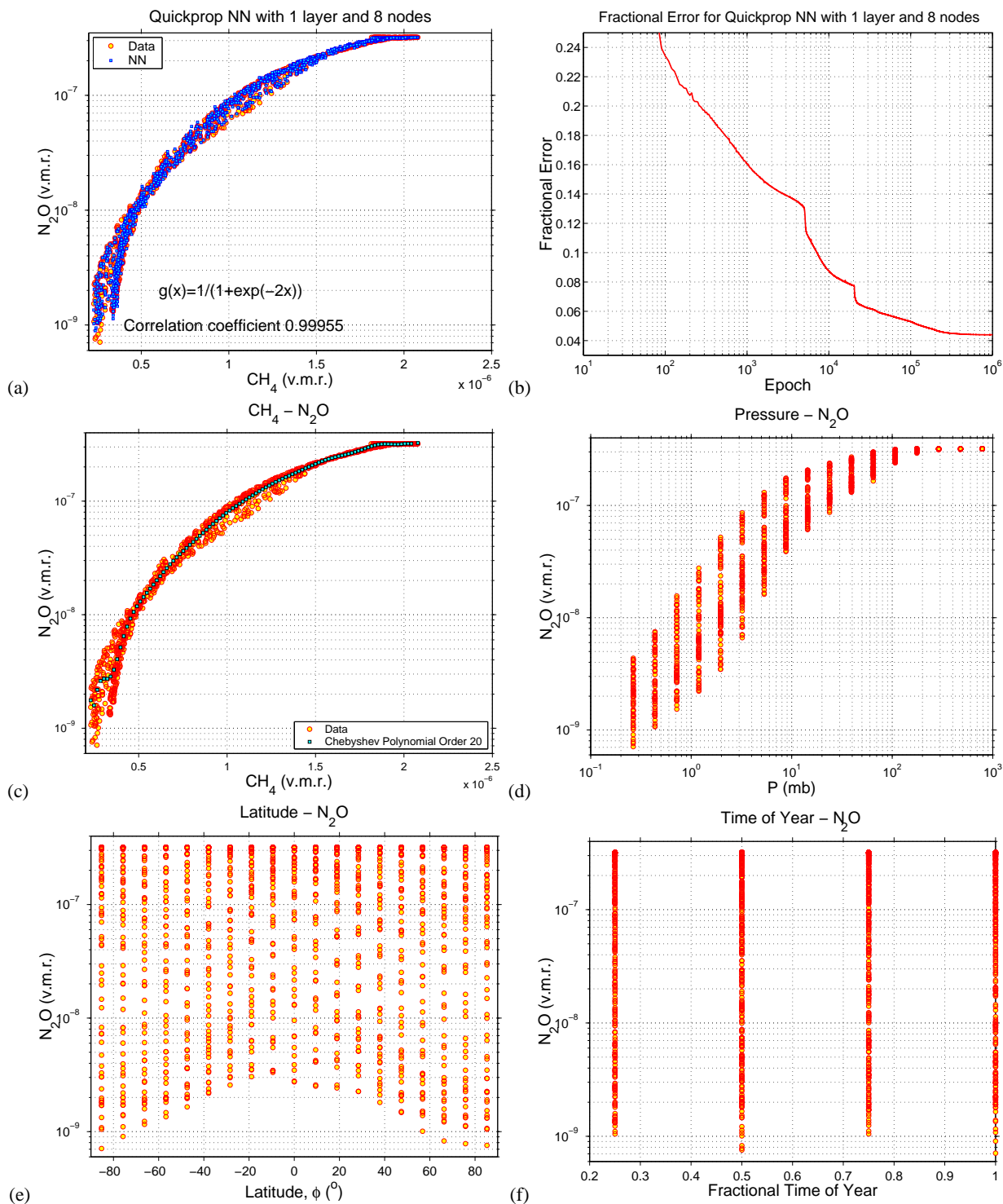


Fig. 1. The neural network used to produce the CH_4 - N_2O correlation in Panel (a) used Quickprop learning and one hidden layer with eight nodes. The correlation coefficient between the actual solution and the neural network solution was 0.9995. Panel (b) shows how the median fractional error of the neural network decreases with epoch (iteration). Both CH_4 and pressure are strongly correlated with N_2O as can be seen in panels (c) and (d). Latitude and time are only weakly correlated with N_2O as can be seen in panels (e) and (f). Even though the correlation with time of year and latitude is relatively weak it still does play a role in capturing some of the details of the CH_4 - N_2O correlation in Panel (a).

2 Neural networks

Computational neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between elements (Fig. 1b).

Commonly neural networks are 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 used, in this supervised learning, to train a network. Batch training of a network proceeds by making weight and bias changes based on an entire set (batch) of input vectors. Incremental training changes the weights and biases of a network as needed after presentation of each individual input vector. Incremental training is sometimes referred to as “on line” or “adaptive” training.

Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision and control systems. It is well established that multilayer feed-forward networks are universal approximators (Hornik et al., 1989; Castro and Delgado, 1996; Ying, 1998).

In this study we use neural networks (Peterson et al., 1994) to describe the temporal and spatial dependence of tracer correlations (Fig. 1).

To find the optimum neural network configuration a range of network architectures were considered containing between one and two hidden layers with between one and sixteen nodes in each hidden layer. A range of updating procedures were also used including back-propagation, Manhattan learning, Langevin Learning, Quickprop and Rprop. Each network was trained for 10^6 epochs. The details of the different learning methods can be found in (Peterson et al., 1994). A variety of activation functions were used. Non-linear activation functions performed best, and the most successful is shown below in Eq. (1). To determine which network architecture and updating procedure was most suitable each configuration was tried in turn and the correlation coefficient between the actual solution and the neural network solution were computed (the correlation coefficient being a normalized measure of the linear relationship strength between variables). The configuration with the highest correlation coefficient between simulated and training values was chosen. This configuration used the Quickprop (Qprop) method, introduced by Fahlman (1988), a very popular batch training algorithm for Feedforward Neural Networks.

2.1 The CH₄-N₂O correlation

Figure 1a shows the CH₄-N₂O correlation from the Cambridge 2D model (Law and Pyle, 1993a,b) overlaid with a

neural network fit to the correlation. The neural network used was a feed-forward multilayer perceptron type with Quickprop learning (Peterson et al., 1994). There were four inputs, one output, and one hidden layer with eight nodes. A non-linear activation function was used, namely

$$g(x) = \frac{1}{1 + \exp(-2x)} \quad (1)$$

The training dataset contained 1292 patterns, sampling the input space completely as shown in Fig. 1. The network was constrained for 10^6 epochs (iterations).

The correlation coefficient between the actual solution and the neural network solution was 0.9995. Fig. 1 panel (b) shows how the median fractional error of the neural network decreases with epoch (iteration). Both CH₄ and pressure are strongly correlated with N₂O as can be seen in panels (c) and (d). Latitude and time are only weakly correlated with N₂O as can be seen in panels (e) and (f). Even though the correlation with time of year and latitude is relatively weak it still does play a role in capturing some of the details of the CH₄-N₂O correlation in Panel (a).

A polynomial or other fit will typically do a good job of describing the CH₄-N₂O correlation for high values of CH₄ and N₂O. However, for low values of CH₄ and N₂O there is quite a spread in the relationship which a single curve can not describe. This is the altitude dependent regime where the correlation shows significant variation with altitude (Minschwaner et al., 1996).

Fig. 1c shows a more conventional fit using a Chebyshev polynomial of order 20. This fit was chosen as giving the best agreement to the CH₄-N₂O correlation after performing fits using 3667 different equations. Even though this is a good fit the spread of values can not be described by a single curve. However, a neural network trained with the latitude, pressure, time of year, and CH₄ volume mixing ratio (v.m.r.) (four inputs) is able to well reproduce the N₂O v.m.r. (one output), including the spread for low values of CH₄ and N₂O.

2.2 Scaling

Variable scaling often allows neural networks to achieve better results. In this case all variables were scaled to vary between zero and one. If the initial range of values was more than an order of magnitude then log scaling was also applied. In the case of time of year the sine of the fractional time of year was used to avoid a step discontinuity at the start of the year.

3 Conclusions

Neural networks are ideally suited to describe the spatial and temporal dependence of tracer-tracer correlations. Even in regions when the correlations are less compact. Useful insight can be gained into the relative roles of the input variables from visualizing the network weight assignment.

The neural network Fortran code used is available for download.

Acknowledgements. It is a pleasure to acknowledge: NASA for a distinguished Goddard Fellowship in Earth Science; The Royal Society for a Royal Society University Research Fellowship; The government of Israel for an Alon Fellowship; The NERC, EU, and ESA for research support.

References

- Castro, J. L. and Delgado, M.: Fuzzy systems with defuzzification are universal approximators, *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, 26, 149–152, 1996.
- Elkins, J. W., Fahey, D. W., Gilligan, J. M., Dutton, G. S., Baring, T. J., Volk, C. M., Dunn, R. E., Myers, R. C., Montzka, S. A., Wamsley, P. R., Hayden, A. H., Butler, J. H., Thompson, T. M., Swanson, T. H., Dlugokencky, E. J., Novelli, P. C., Hurst, D. F., Lobert, J. M., Ciciora, S. J., McLaughlin, R. J., Thompson, T. L., Winkler, R. H., Fraser, P. J., Steele, L. P., and Lucarelli, M. P.: Airborne gas chromatograph for in situ measurements of long-lived species in the upper troposphere and lower stratosphere, *Geophys. Res. Lett.*, 23, 347–350, 1996.
- Fahey, D. W., Murphy, D. M., Kelly, K. K., Ko, M. K. W., Proffitt, M. H., Eubank, C. S., Ferry, G. V., Loewenstein, M., and Chan, K. R.: Measurements of nitric-oxide and total reactive nitrogen in the antarctic stratosphere – observations and chemical implications, *J. Geophys. Res. (Atmos.)*, 94, 16 665–16 681, 1989.
- Fahlman, S.: Faster-learning variations on back propagation: an empirical study, in: *Proceedings of the 1988 Connectionist Models Summer School*, edited by Touretzky, D., Hinton, G., and Sejnowski, T., Morgan Kaufmann, San Mateo, CA, pp. 38–51, 1988.
- Fischer, H., Wienhold, F. G., Hoor, P., Bujok, O., Schiller, C., Siegmund, P., Ambaum, M., Scheeren, H. A., and Lelieveld, J.: Tracer correlations in the northern high latitude lowermost stratosphere: Influence of cross-tropopause mass exchange, *Geophys. Res. Lett.*, 27, 97–100, 2000.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359–366, 1989.
- Keim, E. R., Loewenstein, M., Podolske, J. R., Fahey, D. W., Gao, R. S., Woodbridge, E. L., Wamsley, R. C., Donnelly, S. G., Del-Negro, L. A., Nevison, C. D., Solomon, S., Rosenlof, K. H., Scott, C. J., Ko, M. K. W., Weisenstein, D., and Chan, K. R.: Measurements of the NO_y-N₂O correlation in the lower stratosphere: Latitudinal and seasonal changes and model comparisons, *J. Geophys. Res. (Atmos.)*, 102, 13 193–13 212, 1997.
- Law, K. and Pyle, J.: Modeling trace gas budgets in the troposphere, 1. O₃ and odd nitrogen, *J. Geophys. Res. (Atmos.)*, 98, 18 377–18 400, 1993a.
- Law, K. and Pyle, J.: Modeling trace gas budgets in the troposphere, 2. CH₄ and CO, *J. Geophys. Res. (Atmos.)*, 98, 18 401–18 412, 1993b.
- Loewenstein, M., Podolske, J. R., Fahey, D. W., Woodbridge, E. L., Tin, P., Weaver, A., Newman, P. A., Strahan, S. E., Kawa, S. R., Schoeberl, M. R., and Lait, L. R.: New observations of the NO_y/N₂O correlation in the lower stratosphere, *Geophys. Res. Lett.*, 20, 2531–2534, 1993.
- Michelson, H. A., Manney, G. L., Gunson, M. R., and Zander, R.: Correlations of stratospheric abundances of NO_y, O₃, N₂O, and CH₄ derived from ATMOS measurements, *J. Geophys. Res. (Atmos.)*, 103, 28 347–28 359, 1998.
- Minschwaner, K., Dessler, A. E., Elkins, J. W., Volk, C. M., Fahey, D. W., Loewenstein, M., Podolske, J. R., Roche, A. E., and Chan, K. R.: Bulk properties of isentropic mixing into the tropics in the lower stratosphere, *J. Geophys. Res. (Atmos.)*, 101, 9433–9439, 1996.
- Morgenstern, O., Lee, A. M., and Pyle, J. A.: Cumulative mixing inferred from stratospheric tracer relationships, *J. Geophys. Res. (Atmos.)*, 108, 2002.
- Muscari, G., de Zafra, R. L., and Smyshlyaev, S.: Evolution of the NO_y-N₂O correlation in the antarctic stratosphere during 1993 and 1995, *J. Geophys. Res. (Atmos.)*, 108, 2003.
- Peterson, C., Rognvaldsson, T., and Lonnblad, L.: JETNET 3.0 a versatile artificial neural network package, *Comp. Phys. Comm.*, 81, 185–220, 1994.
- Plumb, R. A. and Ko, M. K. W.: Interrelationships between mixing ratios of long lived stratospheric constituents, *J. Geophys. Res. (Atmos.)*, 97, 10 145–10 156, 1992.
- Reber, C. A., Trevathan, C. E., McNeal, R. J., and Luther, M. R.: The upper-atmosphere research satellite (uars) mission, *J. Geophys. Res. (Atmos.)*, 98, 10 643–10 647, 1993.
- Rinsland, C. P., Salawitch, R. J., Gunson, M. R., Solomon, S., Zander, R., Mahieu, E., Goldman, A., Newchurch, M. J., Irion, F. W., and Chang, A. Y.: Polar stratospheric descent of NO_y and CO and arctic denitrification during winter 1992–1993, *J. Geophys. Res. (Atmos.)*, 104, 1847–1861, 1999.
- Russell, J. M., Gordley, L. L., Park, J. H., Drayson, S. R., Hesketh, W. D., Cicerone, R. J., Tuck, A. F., Frederick, J. E., Harries, J. E., and Crutzen, P. J.: The halogen occultation experiment, *J. Geophys. Res. (Atmos.)*, 98, 10 777–10 797, 1993.
- Schoeberl, M. R., Roche, A. E., Russell, J. M., Ortland, D., Hays, P. B., and Waters, J. W.: An estimation of the dynamical isolation of the tropical lower stratosphere using UARS wind and trace gas observations of the quasibiennial oscillation, *Geophys. Res. Lett.*, 24, 53–56, 1997.
- Strahan, S. E.: Climatologies of lower stratospheric NO_y and O₃ and correlations with N₂O based on in situ observations, *J. Geophys. Res. (Atmos.)*, 104, 30 463–30 480, 1999.
- Wahg, D. W. and Funatsu, B. M.: Intrusions into the tropical upper troposphere: Three-dimensional structure and accompanying ozone and olr distributions, *J. Atmos. Sci.*, 60, 637–653, 2003.
- Ying, H.: General takagi-sugeno fuzzy systems with simplified linear rule consequent are universal controllers, models and filters, *Information Sciences*, 108, 91–107, 1998.