We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Artificial Intelligence in Geoscience and Remote Sensing

David John Lary
Joint Center for Earth Systems Technology (JCET) UMBC, NASA/GSFC
United States

1. Introduction

Machine learning has recently found many applications in the geosciences and remote sensing. These applications range from bias correction to retrieval algorithms, from code acceleration to detection of disease in crops. As a broad subfield of artificial intelligence, machine learning is concerned with algorithms and techniques that allow computers to "learn". The major focus of machine learning is to extract information from data automatically by computational and statistical methods.

Over the last decade there has been considerable progress in developing a machine learning methodology for a variety of Earth Science applications involving trace gases, retrievals, aerosol products, land surface products, vegetation indices, and most recently, ocean products (*Yi and Prybutok*, 1996, *Atkinson and Tatnall*, 1997, *Carpenter et al.*, 1997, *Comrie*, 1997, *Chevallier et al.*, 1998, *Hyyppa et al.*, 1998, *Gardner and Dorling*, 1999, *Lary et al.*, 2004, *Lary et al.*, 2007, *Brown et al.*, 2008, *Lary and Aulov*, 2008, *Caselli et al.*, 2009, *Lary et al.*, 2009). Some of this work has even received special recognition as a NASA Aura Science highlight (*Lary et al.*, 2007) and commendation from the NASA MODIS instrument team (*Lary et al.*, 2009). The two types of machine learning algorithms typically used are neural networks and support vector machines. In this chapter, we will review some examples of how machine learning is useful for Geoscience and remote sensing, these examples come from the author's own research.

2. Typical Applications

One of the features that make machine-learning algorithms so useful is that they are "universal approximators". They can learn the behaviour of a system if they are given a comprehensive set of examples in a training dataset. These examples should span as much of the parameter space as possible. Effective learning of the system's behaviour can be achieved even if it is multivariate and non-linear. An additional useful feature is that we do not need to know a priori the functional form of the system as required by traditional least-squares fitting, in other words they are non-parametric, non-linear and multivariate learning algorithms.

The uses of machine learning to date have fallen into three basic categories which are widely applicable across all of the Geosciences and remote sensing, the first two categories use machine learning for its regression capabilities, the third category uses machine learning for its

classification capabilities. We can characterize the three application themes are as follows: First, where we have a theoretical description of the system in the form of a deterministic model, but the model is computationally *expensive*. In this situation, a machine-learning "wrapper" can be applied to the deterministic model providing us with a "code accelerator". A good example of this is in the case of atmospheric photochemistry where we need to solve a large coupled system of ordinary differential equations (ODEs) at a large grid of locations. It was found that applying a neural network wrapper to the system was able to provide a speed up of between a factor of 2 and 200 depending on the conditions. Second, when we do not have a deterministic model but we have data available enabling us to empirically learn the behaviour of the system. Examples of this would include: Learning inter-instrument bias between sensors with a temporal overlap, and inferring physical parameters from remotely sensed proxies. Third, machine learning can be used for classification, for example, in providing land surface type classifications. Support Vector Machines perform particularly well for classification problems.

Now that we have an overview of the typical applications, the sections that follow will introduce two of the most powerful machine learning approaches, neural networks and support vector machines and then present a variety of examples.

3. Machine Learning

3.1 Neural Networks

Neural networks are multivariate, non-parametric, 'learning' algorithms (*Haykin*, 1994, *Bishop*, 1995, 1998, *Haykin*, 2001a, *Haykin*, 2001b, 2007) inspired by biological neural networks. Computational neural networks (NN) consist of an interconnected group of artificial neurons that processes information in parallel using a connectionist approach to computation. A NN is a non-linear statistical data-modelling tool that can be used to model complex relationships between inputs and outputs or to find patterns in data. The basic computational element of a NN is a model neuron or node. A node receives input from other nodes, or an external source (e.g. the input variables). A schematic of an example NN is shown in Figure 1. Each input has an associated weight, *w*, that can be modified to mimic synaptic learning. The unit computes some function, *f,* of the weighted sum of its inputs:

$$y_i = f\left(\sum_j w_{ij} y_j\right)$$

Its output, in turn, can serve as input to other units. w_{ij} refers to the weight from unit *j* to unit *i*. The function *f* is the node's activation or transfer function. The transfer function of a node defines the output of that node given an input or set of inputs. In the simplest case, *f* is the identity function, and the unit's output is y_i, this is called a linear node. However, non-linear sigmoid functions are often used, such as the hyperbolic tangent sigmoid transfer function and the log-sigmoid transfer function. Figure 1 shows an example feed-forward perceptron NN with five inputs, a single output, and twelve nodes in a hidden layer. A perceptron is a computer model devised to represent or simulate the ability of the brain to recognize and discriminate. In most cases, a NN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase.
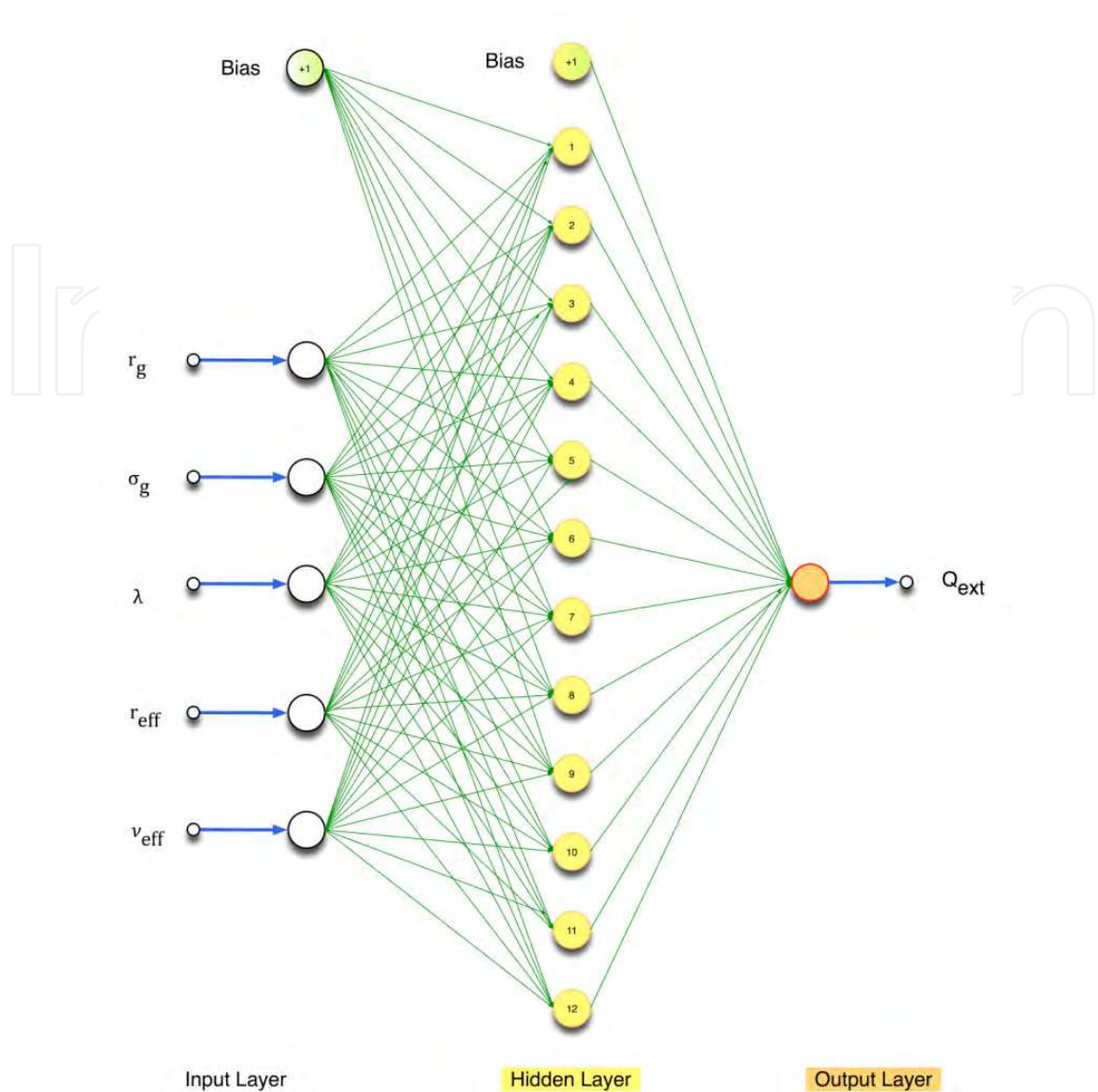
Fig. 1. Example neural network architecture showing a network with five inputs, one output, and twelve hidden nodes.

When we perform neural network training, we want to ensure we can independently assess the quality of the machine learning 'fit'. To insure this objective assessment we usually randomly split our training dataset into three portions, typically of 80%, 10% and 10%. The largest portion containing 80% of the dataset is used for training the neural network weights. This training is iterative, and on each training iteration we evaluate the current root mean square (RMS) error of the neural network output. The RMS error is calculated by using the second 10% portion of the data that was not used in the training. We use the RMS error and the way the RMS error changes with training iteration (epoch) to determine the convergence of our training. When the training is complete, we then use the final 10% portion of data as a totally independent validation dataset. This final 10% portion of the data is randomly chosen from the training dataset and is not used in either the training or RMS evaluation. We only use the neural network if the validation scatter diagram, which plots the actual data from validation portion against the neural network estimate, yields a straight-line graph with a

slope very close to one and an intercept very close to zero. This is a stringent, independent and objective validation metric. The validation is global as the data is randomly selected over all data points available. For our studies, we typically used feed-forward back-propagation neural networks with a Levenberg-Marquardt back-propagation training algorithm (*Levenberg*, 1944, *Marquardt*, 1963, *Moré*, 1977, *Marquardt*, 1979).

3.2 Support Vector Machines

Support Vector Machines (SVM) are based on the concept of decision planes that define decision boundaries and were first introduced by Vapnik (*Vapnik*, 1995, 1998, 2000) and has subsequently been extended by others (*Scholkopf et al.*, 2000, *Smola and Scholkopf*, 2004). A decision plane is one that separates between a set of objects having different class memberships. The simplest example is a linear classifier, i.e. a classifier that separates a set of objects into their respective groups with a line. However, most classification tasks are not that simple, and often more complex structures are needed in order to make an optimal separation, i.e., correctly classify new objects (test cases) on the basis of the examples that are available (training cases). Classification tasks based on drawing separating lines to distinguish between objects of different class memberships are known as hyperplane classifiers.

SVMs are a set of related supervised learning methods used for classification and regression. Viewing input data as two sets of vectors in an n-dimensional space, an SVM will construct a separating hyperplane in that space, one that maximizes the margin between the two data sets. To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the neighboring data points of both classes, since in general the larger the margin the better the generalization error of the classifier. We typically used the SVMs provided by LIBSVM (*Fan et al.*, 2005, *Chen et al.*, 2006).

4. Applications

Let us now consider some applications.

4.1 Bias Correction: Atmospheric Chlorine Loading for Ozone Hole Research

Critical in determining the speed at which the stratospheric ozone hole recovers is the total amount of atmospheric chlorine. Attributing changes in stratospheric ozone to changes in chlorine requires knowledge of the stratospheric chlorine abundance over time. Such attribution is central to international ozone assessments, such as those produced by the World Meteorological Organization (*Wmo*, 2006). However, we do not have continuous observations of all the key chlorine gases to provide such a continuous time series of stratospheric chlorine. To address this major limitation, we have devised a new technique that uses the long time series of available hydrochloric acid observations and neural networks to estimate the stratospheric chlorine (Cl_y) abundance (*Lary et al.*, 2007).

Knowledge of the distribution of inorganic chlorine Cl_y in the stratosphere is needed to attribute changes in stratospheric ozone to changes in halogens, and to assess the realism of chemistry-climate models (*Eyring et al.*, 2006, *Eyring et al.*, 2007, *Waugh and Eyring*, 2008). However, simultaneous measurements of the major inorganic chlorine species are rare (*Zander et al.*, 1992, *Gunson et al.*, 1994, *Webster et al.*, 1994, *Michelsen et al.*, 1996, *Rinsland et al.*, 1996,

Zander et al., 1996, *Sen et al.*, 1999, *Bonne et al.*, 2000, *Voss et al.*, 2001, *Dufour et al.*, 2006, *Nassar et al.*, 2006). In the upper stratosphere, the situation is a little easier as Cl_y can be inferred from HCl alone (e.g., (*Anderson et al.*, 2000, *Froidevaux et al.*, 2006b, *Santee et al.*, 2008)). Our new estimates of stratospheric chlorine using machine learning (*Lary et al.*, 2007) work throughout the stratosphere and provide a much-needed critical test for current global models. This critical evaluation is necessary as there are significant differences in both the stratospheric chlorine and the timing of ozone recovery in the available model predictions.

Hydrochloric acid is the major reactive chlorine gas throughout much of the atmosphere, and throughout much of the year. However, the observations of HCl that we do have (from UARS HALOE, ATMOS, SCISAT-1 ACE and Aura MLS) have significant biases relative to each other. We found that machine learning can also address the inter-instrument bias (*Lary et al.*, 2007, *Lary and Aulov*, 2008). We compared measurements of HCl from the different instruments listed in Table 1. The Halogen Occultation Experiment (HALOE) provides the longest record of space based HCl observations. Figure 2 compares HALOE HCl with HCl observations from (a) the Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS), (b) the Atmospheric Chemistry Experiment (ACE) and (c) the Microwave Limb Sounder (MLS).
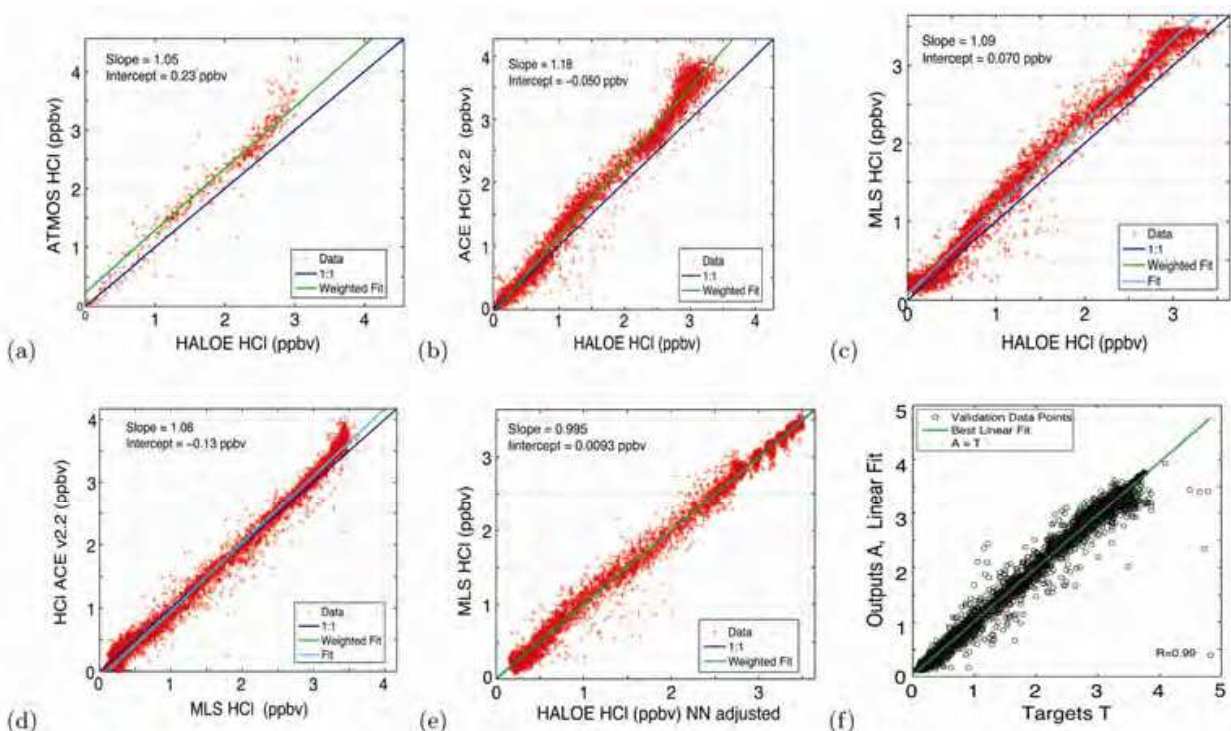

Fig. 2. Panels (a) to (d) show scatter plots of all contemporaneous observations of HCl made by HALOE, ATMOS, ACE and MLS Aura. In panels (a) to (c) HALOE is shown on the x-axis. Panel (e) correspond to panel (c) except that it uses the neural network 'adjusted' HALOE HCl values. Panel (f) shows the validation scatter diagram of the neural network estimate of $Cl_y \approx HCl + ClONO_2 + ClO + HOCl$ versus the actual Cl_y for a totally independent data sample *not* used in training the neural network.

A consistent picture is seen in these plots: HALOE HCl measurements are lower than those from the other instruments. The slopes of the linear fits (relative scaling) are 1.05 for the HALOE-ATMOS comparison, 1.09 for the HALOE-MLS, and 1.18 for the HALOE-ACE. The

offsets are apparent at the 525 K isentropic surface and above. Previous comparisons among HCl datasets reveal a similar bias for HALOE (*Russell et al.*, 1996, *Mchugh et al.*, 2005, *Froidevaux et al.*, 2006a, *Froidevaux et al.*, 2008). ACE and MLS HCl measurements are in much better agreement (Figure 2d). Note, the measurements agree within the stated observational uncertainties summarized in Table 1.

Instrument	Temporal Coverage	Species	References	Median Observation Uncertainty
ACE v2.2	2004-2006	HCl, ClONO$_2$ ClO, and HOCl	*Bernath et al.* [2005]	8% (HCl), 30% (ClONO$_2$) >100% (ClO), >100% (HOCl)
ATMOS	1991, 1993, 1994	HCl, ClONO$_2$	*Zander et al.* [1992]	8% (HCl), 60% (ClONO$_2$)
Aura MLS v1	2004-2006	HCl, ClO and HOCl	*Froidevaux et al.* [2006a]	12% (HCl), 76% (ClO), >100% (HOCl)
CLAES v9	1991-1993	ClONO$_2$	*Roche et al.* [1993]	>100%
CRISTA	1994, 1997	ClONO$_2$	*Offermann et al.* [1999]	61%
HALOE v19	1991-2005	HCl	*Russell et al.* [1993]	4%
UARS MLS v5	1991-1999	ClO	*Waters et al.* [1996]	>100% (ClO)

Table 1. The instruments and constituents used in constructing the Cl$_y$ record from 1991-2006. The uncertainties given are the median values calculated for each level 2 measurement profile and its uncertainty (both in mixing ratio) for all the observations made. The uncertainties are larger than usually quoted for MLS ClO because they reflect the single profile precision, which is improved by temporal and/or spatial averaging. The HALOE uncertainties are only estimates of random error and do not include any indications of overall accuracy.

To combine the above HCl measurements to form a continuous time series of HCl (and then Cl$_y$) from 1991 to 2006 it is necessary to account for the biases between data sets. A neural network is used to learn the mapping from one set of measurements onto another as a function of equivalent latitude and potential temperature. We consider two cases. In one case ACE HCl is taken as the reference and the HALOE and Aura HCl observations are adjusted to agree with ACE HCl. In the other case HALOE HCl is taken as the reference and the Aura and ACE HCl observations are adjusted to agree with HALOE HCl. In both cases we use equivalent latitude and potential temperature to produce average profiles. The purpose of the NN mapping is simply to learn the bias as a function of location, not to imply which instrument is correct.

The precision of the correction using the neural network mapping is of the order of ±0.3 ppbv, as seen in Figure 2 (e) that shows the results when HALOE HCl measurements have been mapped into ACE measurements. The mapping has removed the bias between the measurements and has straightened out the 'wiggles' in 2 (c), i.e., the neural network has learned the equivalent PV latitude and potential temperature dependence of the bias between HALOE and MLS. The inter-instrument offsets are not constant in space or time, and are not a simple function of Cl$_y$.

So employing neural networks allows us to: Form a seamless record of HCl using observations from several space-borne instruments using neural networks. Provide an estimated of the associated inter-instrument bias. Infer Cl$_y$ from HCl, and thereby provide a seamless record of Cl$_y$, the parameter needed for examining the ozone hole recovery. A similar use of machine learning has been made for Aerosol Optical Depths, the subject of the next sub-section.
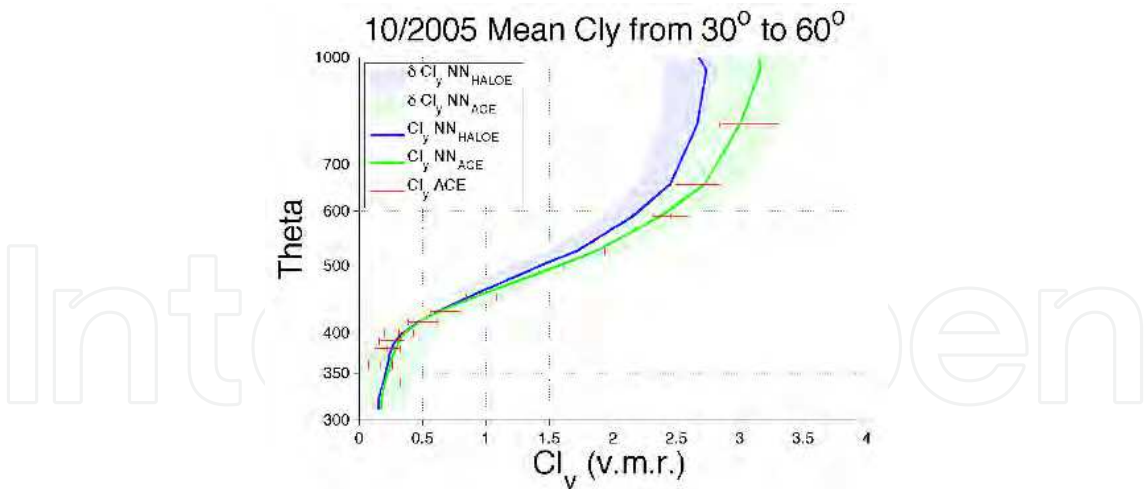
Fig. 3. Cl_y average profiles between 30° and 60°N for October 2005, estimated by neural network calibrated to HALOE HCl (blue curve), estimated by neural network calibrated to ACE HCl (green), or from ACE observations of HCl, $ClONO_2$, ClO, and HOCl (red crosses). In each case, the shaded range represents the total uncertainty; it includes the observational uncertainty, the representativeness uncertainty (the variability over the analysis grid cell), the neural network uncertainty. The vertical extent of this plot was limited to below 1000 K (≈35 km), as there is no ACE v2.2 ClO data for the upper altitudes. In addition, above ≈750 K (≈25 km), ClO constitutes a larger fraction of Cl_y (up to about 10%) and so the large uncertainties in ClO have greater effect.
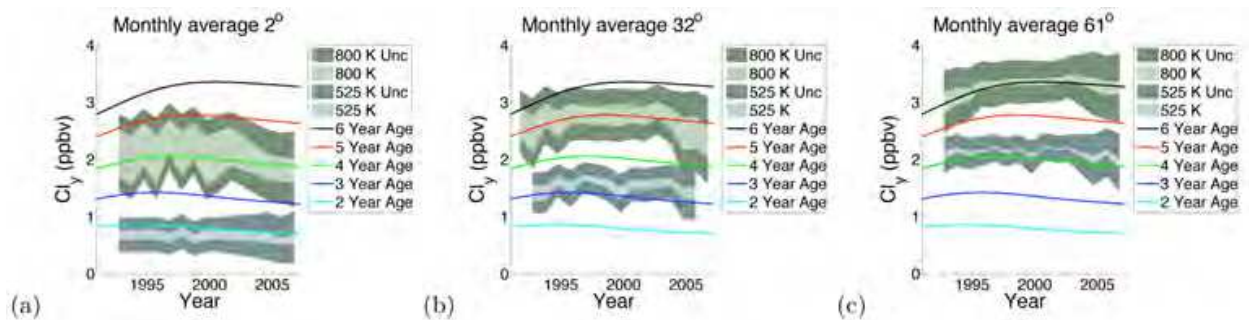

Fig. 4. Panels (a) to (c) show October Cl_y time-series for the 525 K isentropic surface (≈20 km) and the 800 K isentropic surface (≈30 km). In each case the dark shaded range represents the total uncertainty in our estimate of Cl_y. This total uncertainty includes the observational uncertainty, the representativeness uncertainty (the variability over the analysis grid cell), the inter-instrument bias in HCl, the uncertainty associated with the neural network inter-instrument correction, and the uncertainty associated with the neural network inference of Cl_y from HCl and CH_4. The inner light shading depicts the uncertainty on Cl_y due to the inter-instrument bias in HCl alone. The upper limit of the light shaded range corresponds to the estimate of Cl_y based on all the HCl observations calibrated by a neural network to agree with ACE v2.2 HCl. The lower limit of the light shaded range corresponds to the estimate of Cl_y based on all the HCl observations calibrated to agree with HALOE v19 HCl. Overlaid are lines showing the Cl_y based on age of air calculations (*Newman et al.*, 2006). To minimize variations due to differing data coverage months with less than 100 observations of HCl in the equivalent latitude bin were left out of the time-series.
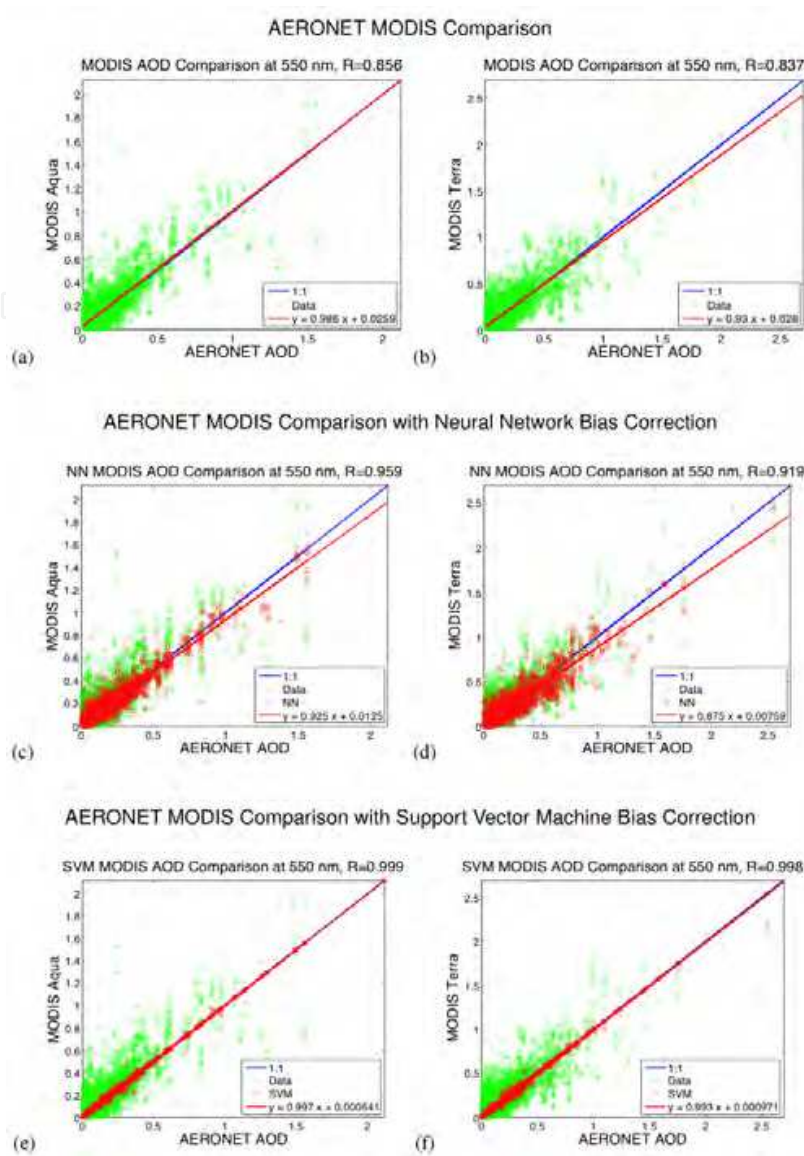
Fig. 5. Scatter diagram comparisons of Aerosol Optical Depth (AOD) from AERONET (x-axis) and MODIS (y-axis) as green circles overlaid with the ideal case of perfect agreement (blue line). The measurements shown in the comparison were made within half an hour of each other, with a great circle separation of less than 0.25° and with a solar zenith angle difference of less than 0.1°. The left hand column of plots is for MODIS Aqua and the right hand column of plots is for MODIS Terra. The first row shows the comparisons between AERONET and MODIS for the entire period of overlap between the MODIS and AERONET instruments from the launch of the MODIS instrument to the present. The second row shows the same comparison overlaid with the neural network correction as red circles. We note that the neural network bias correction makes a substantial improvement in the correlation coefficient with AERONET. An improvement from 0.86 to 0.96 for MODIS Aqua and an improvement from 0.84 to 0.92 for MODIS Terra. The third row shows the comparison overlaid with the support vector regression correction as red circles. We note that the support vector regression bias correction makes an even greater improvement in the correlation coefficient than the neural network correction. An improvement from 0.86 to 0.99 for MODIS Aqua and an improvement from 0.84 to 0.99 for MODIS Terra.

4.2 Bias Correction: Aerosol Optical Depth

As highlighted in the 2007 IPCC report on Climate Change, aerosol and cloud radiative effects remain the largest uncertainties in our understanding of climate change (*Solomon et al.*, 2007). Over the past decade observations and retrievals of aerosol characteristics have been conducted from space-based sensors, from airborne instruments and from ground-based samplers and radiometers. Much effort has been directed at these data sets to collocate observations and retrievals, and to compare results. Ideally, when two instruments measure the same aerosol characteristic at the same time, the results should agree within well-understood measurement uncertainties. When inter-instrument biases exist, we would like to explain them theoretically from first principles. One example of this is the comparison between the aerosol optical depth (AOD) retrieved by the Moderate Resolution Imaging Spectroradiometer (MODIS) and the AOD measured by the Aerosol Robotics Network (AERONET). While progress has been made in understanding the biases between these two data sets, we still have an imperfect understanding of the root causes. (*Lary et al.*, 2009) examined the efficacy of empirical machine learning algorithms for aerosol bias correction.

Machine learning approaches (Neural Networks and Support Vector Machines) were used by (*Lary et al.*, 2009) to explore the reasons for a persistent bias between aerosol optical depth (AOD) retrieved from the MODerate resolution Imaging Spectroradiometer (MODIS) and the accurate ground-based Aerosol Robotics Network (AERONET). While this bias falls within the expected uncertainty of the MODIS algorithms, there is still room for algorithm improvement. The results of the machine learning approaches suggest a link between the MODIS AOD biases and surface type. From figure 5 we can see that machine learning algorithms were able to effectively adjust the AOD bias seen between the MODIS instruments and AERONET. Support vector machines performed the best improving the correlation coefficient between the AERONET AOD and the MODIS AOD from 0.86 to 0.99 for MODIS Aqua, and from 0.84 to 0.99 for MODIS Terra.

Key in allowing the machine learning algorithms to 'correct' the MODIS bias was provision of the surface type and other ancillary variables that explain the variance between MODIS and AERONET AOD. The provision of the ancillary variables that can explain the variance in the dataset is the key ingredient for the effective use of machine learning for bias correction. A similar use of machine learning has been made for vegetation indices, the subject of the next sub-section.

4.3 Bias Correction: Vegetation Indices

Consistent, long term vegetation data records are critical for analysis of the impact of global change on terrestrial ecosystems. Continuous observations of terrestrial ecosystems through time are necessary to document changes in magnitude or variability in an ecosystem (*Tucker et al.*, 2001, *Eklundh and Olsson*, 2003, *Slayback et al.*, 2003). Satellite remote sensing has been the primary way that scientists have measured global trends in vegetation, as the measurements are both global and temporally frequent. In order to extend measurements through time, multiple sensors with different design and resolution must be used together in the same time series. This presents significant problems as sensor band placement, spectral response, processing, and atmospheric correction of the observations can vary significantly and impact the comparability of the measurements (*Brown et al.*, 2006). Even without differences in atmospheric correction, vegetation index values for the same target recorded under identical

conditions will not be directly comparable because input reflectance values differ from sensor to sensor due to differences in sensor design (*Teillet et al.*, 1997, *Miura et al.*, 2006).

Several approaches have previously been taken to integrate data from multiple sensors. (*Steven et al.*, 2003), for example, simulated the spectral response from multiple instruments and with simple linear equations created conversion coefficients to transform NDVI data from one sensor to another. Their analysis is based on the observation that the vegetation index is critically dependent on the spectral response functions of the instrument used to calculate it. The conversion formulas the paper presents cannot be applied to maximum value NDVI datasets because the weighting coefficients are land cover and dataset dependent, reducing their efficacy in mixed pixel situations (*Steven et al.*, 2003). (*Trishchenko et al.*, 2002) created a series of quadratic functions to correct for differences in the reflectance and NDVI to NOAA-9 AVHRR-equivalents (*Trishchenko et al.*, 2002). Both the (*Steven et al.*, 2003) and the (*Trishchenko et al.*, 2002) approaches are land cover and dataset dependent and thus cannot be used on global datasets where multiple land covers are represented by one pixel. (*Miura et al.*, 2006) used hyper-spectral data to investigate the effect of different spectral response characteristics between MODIS and AVHRR instruments on both the reflectance and NDVI data, showing that the precise characteristics of the spectral response had a large effect on the resulting vegetation index. The complex patterns and dependencies on spectral band functions were both land cover dependent and strongly non-linear, thus we see that an exploration of a non-linear approach may be fruitful.

(*Brown et al.*, 2008) experimented with powerful, non-linear neural networks to identify and remove differences in sensor design and variable atmospheric contamination from the AVHRR NDVI record in order to match the range and variance of MODIS NDVI without removing the desired signal representing the underlying vegetation dynamics. Neural networks are 'data transformers' (*Atkinson and Tatnall*, 1997), where the objective is to associate the elements of one set of data to the elements in another. Relationships between the two datasets can be complex and the two datasets may have different statistical distributions. In addition, neural networks incorporate a priori knowledge and realistic physical constraints into the analysis, enabling a transformation from one dataset into another through a set of weighting functions (*Atkinson and Tatnall*, 1997). This transformation incorporates additional input data that may account for differences between the two datasets.

The objective of (*Brown et al.*, 2008) was to demonstrate the viability of neural networks as a tool to produce a long term dataset based on AVHRR NDVI that has the data range and statistical distribution of MODIS NDVI. Previous work has shown that the relationship between AVHRR and MODIS NDVI is complex and nonlinear (*Gallo et al.*, 2003, *Brown et al.*, 2006, *Miura et al.*, 2006), thus this problem is well suited to neural networks if appropriate inputs can be found. The influence of the variation of atmospheric contamination of the AVHRR data through time was explored by using observed atmospheric water vapor from the Total Ozone Mapping Spectrometer (TOMS) instrument during the overlap period 2000-2004 and back to 1985. Examination of the resulting MODIS fitted AVHRR dataset both during the overlap period and in the historical dataset will enable an evaluation of the efficacy of the neural net approach compared to other approaches to merge multiple-sensor NDVI datasets.
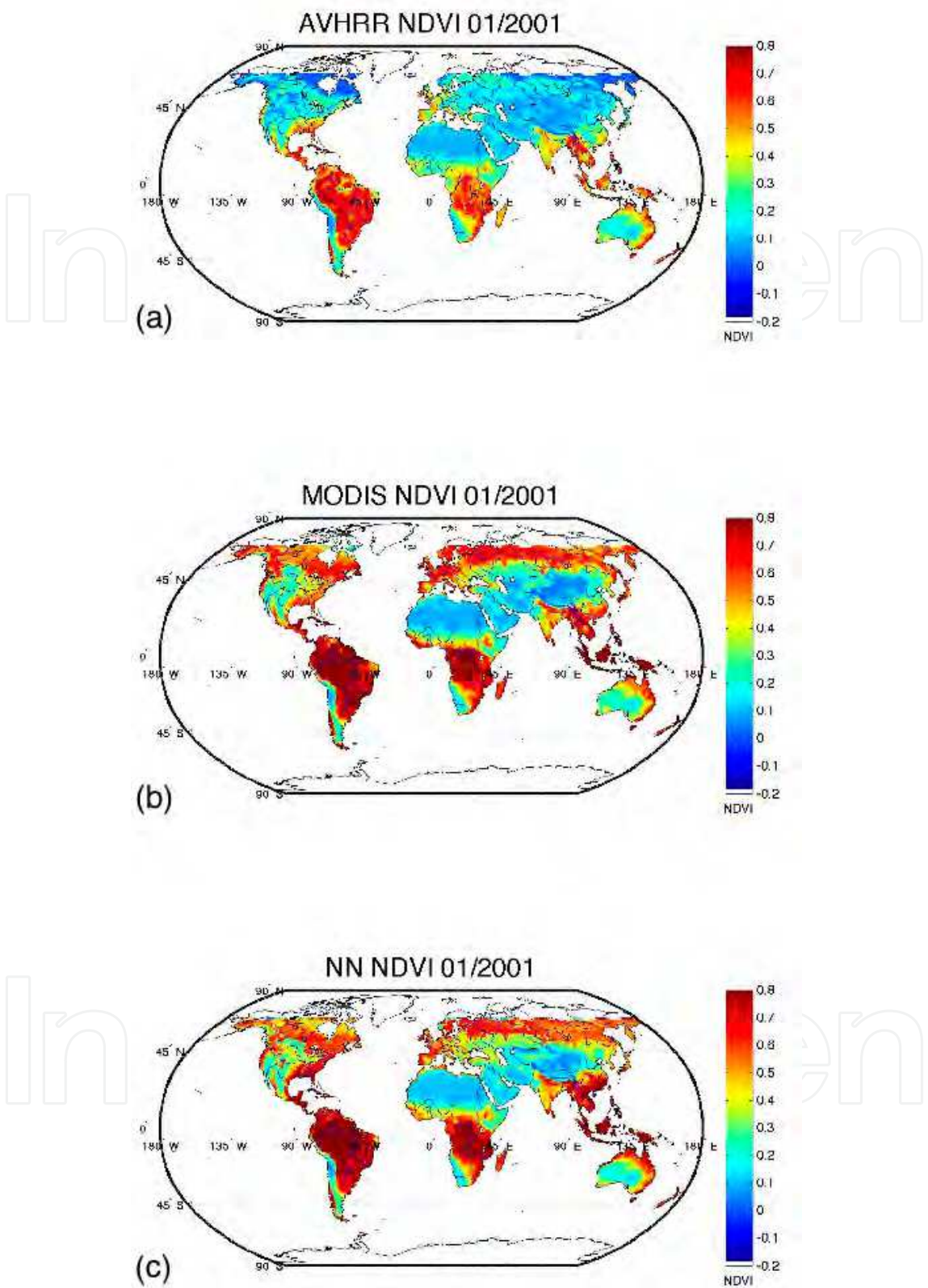
Fig. 6. A comparison of the NDVI from AVHR (panel a), MODIS (panel p), and then a reconstruction of MODIS using AVHRR and machine learning (panel c). We note that the machine learning can successfully account for the large differences that are found between AVHRR and MODIS.

Remote sensing datasets are the result of a complex interaction between the design of a sensor, the spectral response function, stability in orbit, the processing of the raw data, compositing schemes, and post-processing corrections for various atmospheric effects including clouds and aerosols. The interaction between these various elements is often non-linear and non-additive, where some elements increase the vegetation signal to noise ratio (compositing, for example) and others reduce it (clouds and volcanic aerosols) (*Los*, 1998). Thus, although other authors have used simulated data to explore the relationship between AVHRR and MODIS (*Trishchenko et al.*, 2002, *Van Leeuwen et al.*, 2006), these techniques are not directly useful in producing a sensor-independent vegetation dataset that can be used by data users in the near term.
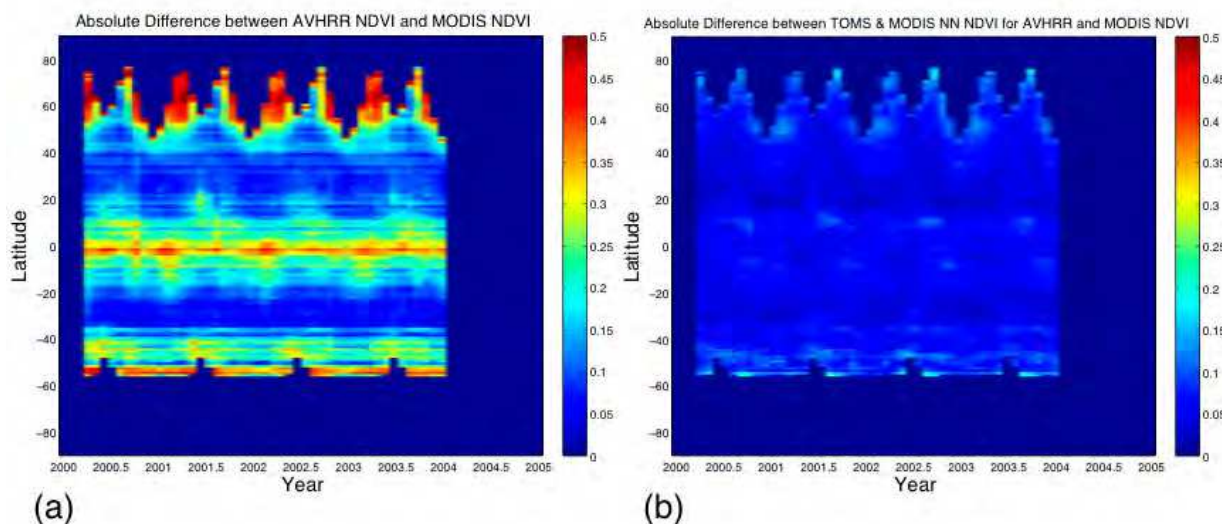

Fig. 7. Panel (a) shows a time-series from 2000 to 2003 of the zonal mean (averaged per latitude) difference between the AVHRR and MODIS NDVIs, this highlights that significant differences exist between the two data products. Panel (b) shows a time series over the same period after the machine learning has been used to "cross-calibrate" AVHRR as MODIS, showing that the machine learning has effectively learnt how to cross-calibrate the instruments.

There are substantial differences between the processed vegetation data from AVHRR and MODIS. (*Brown et al.*, 2008) showed that neural networks are an effective way to have a long data record that utilizes all available data back to 1981 by providing a practical way of incorporating the AVHRR data into a continuum of observations that include both MODIS and VIIRS. The results (*Brown et al.*, 2008) showed that the TOMS data record on clouds, ozone and aerosols can be used to identify and remove sensor-specific atmospheric contaminants that differentially affect the AVHRR over MODIS. Other sensor-related effects, particularly those of changing BRDF, viewing angle, illumination, and other effects that are not accounted for here, remain important sources of additional variability. Although this analysis has not produced a dataset with identical properties to MODIS, it has demonstrated that a neural net approach can remove most of the atmospheric-related aspects of the differences between the sensors, and match the mean, standard deviation and range of the two sensors. A similar technique can be used for the VIIRS sensor once the data is released.

Figure 6 shows a comparison of the NDVI from AVHR (panel a), MODIS (panel p), and then a reconstruction of MODIS using AVHRR and machine learning (panel c). Figure 7 (a) shows a time-series from 2000 to 2003 of the zonal mean difference between the AVHRR and MODIS

NDVIs, this highlights that significant differences exist between the two data products. Panel (b) shows a time series over the same period after the machine learning has been used to "cross-calibrate" AVHRR as MODIS, illustrating that the machine learning has effectively learnt how to cross-calibrate the instruments.

So far, we have seen three examples of using machine learning for bias correction (constituent biases, aerosol optical depth biases and vegetation index biases), and one example of using machine learning to infer a useful proxy from remotely sensed data (Cl_y from HCl). Let us look at one more example of inferring proxies from existing remotely sensed data before moving onto consider using machine learning for code acceleration.

4.4 Inferring Proxies: Tracer Correlations

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. The correlations exist for all long-lived tracers - not just those that 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 using measurements of one tracer to infer the abundance of another tracer. Using these correlations is also as a diagnostic of mixing and can distinguish between air-parcels of different origins. Of special interest are the so-called 'long-lived' tracers: constituents such as nitrous oxide (N_2O), methane (CH_4), 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.

A family of correlations usually achieves the description of such spatially and temporally dependent correlations. However, a single neural network is a natural and effective alternative. The motivation for this case study was preparation for a long-term chemical assimilation of Upper Atmosphere Research Satellite (UARS) 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) but not observations of N_2O 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 CH_4 observations. Machine learning is 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 methane CH_4-N_2O correlation can be well described using a neural network (*Lary et al.*, 2004) trained with the latitude, pressure, time of year, and CH_4 volume mixing ratio (v.m.r.). *Lary et al.* (2004) used a neural network to reproduce the CH_4-N_2O 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. For example, the Halogen Occultation Experiment (HALOE) that continuously observed CH_4 (but not N_2O) from 1991 until 2005.
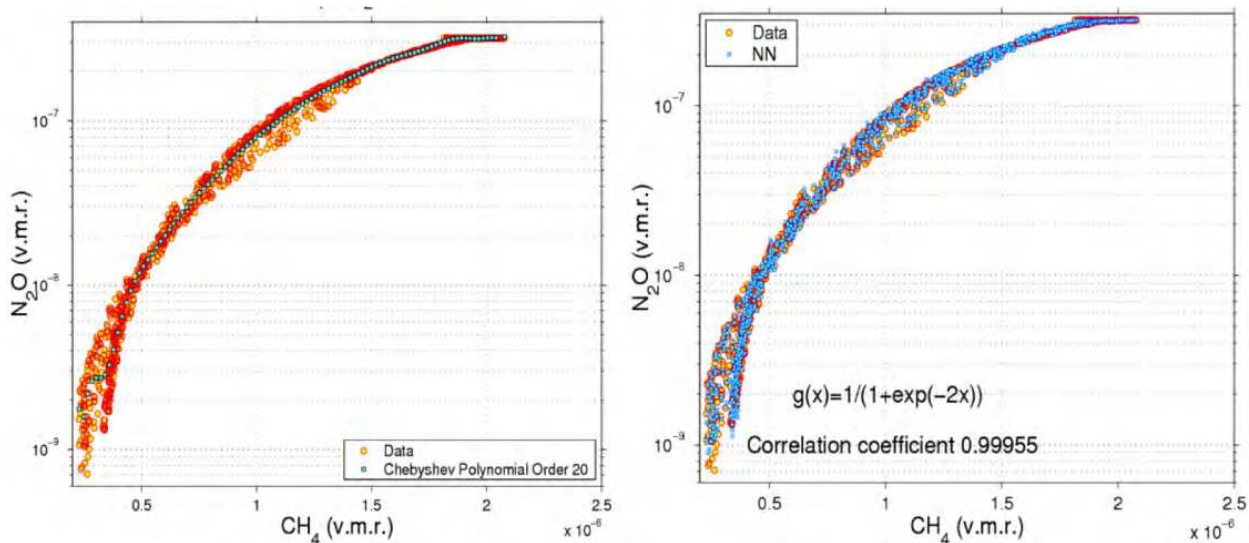
Fig. 8. Panel (a) shows the global N_2O-CH_4 correlation for an entire year, after evaluating the efficacy of 3,000 different functional forms for parametric fits, we overlaid the best, an order 20 Chebyshev Polynomial. However, this still does not account for the multi-variate nature of the problem exhibited by the 'cloud' of points rather than a compact 'curve' or 'line'. However, in panel (b) we can see that a neural network is able to account for the non-linear and multi-variate aspects, the training dataset exhibited a 'cloud' of points, the neural network fit reproduces a 'cloud' of points. The most important factor in producing a 'spread' in the correlations is the strong altitude dependence of the N_2O-CH_4 correlation.

Figure 8 (a) shows the global N_2O-CH_4 correlation for an entire year, after evaluating the efficacy of 3,000 different functional forms for parametric fits, we overlaid the best, an order 20 Chebyshev Polynomial. However, this still does not account for the multi-variate nature of the problem exhibited by the 'cloud' of points rather than a compact 'curve' or 'line'. However, in Figure 8 (b) we can see that a neural network is able to account for the non-linear and multi-variate aspects, the training dataset exhibited a 'cloud' of points, the neural network fit reproduces a 'cloud' of points. The most important factor in producing a 'spread' in the correlations is the strong altitude dependence of the N_2O-CH_4 correlation.

4.5 Code Acceleration: Example from Ordinary Differential Equation Solvers

There are many applications in the Geosciences and remote sensing which are computationally expensive. Machine learning can be very effective in accelerating components of these calculations. We can readily create training datasets for these applications using the very models we would like to accelerate.

The first example for which we found this effective was solving ordinary differential equations. An adequate photochemical mechanism to describe the evolution of ozone in the upper troposphere and lower stratosphere (UT/LS) in a computational model involves a comprehensive treatment of reactive nitrogen, hydrogen, halogens, hydrocarbons, and interactions with aerosols. Describing this complex interaction is computationally expensive, and applications are limited by the computational burden. Simulations are often made tractable by using a coarser horizontal resolution than would be desired or by reducing the interactions accounted for in the photochemical mechanism. These compromises also limit the scientific applications. Machine learning algorithms offer a means to obtain a fast and accurate

solution to the stiff ordinary differential equations that comprise the photochemical calculations, thus making high-resolution simulations including the complete photochemical mechanism much more tractable.

For the sake of an example, a 3D model of atmospheric chemistry and transport, the GMI-COMBO model, can use 55 vertical levels and a 4° latitude x 5° longitude grid and 125 species. With 15-minute time steps the chemical ODE solver is called 119,750,400 times in simulating just one week. If the simulation is for a year then the ODE solver needs to be called 6,227,020,800 (or 6×10^9) times. If the spatial and temporal resolution is doubled then the chemical ODE solver needs to be called a staggering 2.5×10^{10} times to simulate a year. This represents a major computational cost in simulating a constituent's spatial and temporal evolution. The ODEs solved at adjacent grid cells and time steps are very similar. Therefore, if the simulations from one grid cell and time step could be used to speed up the simulation for adjacent grid cells and subsequent time steps, we would have a strategy to dramatically decrease the computational cost of our simulations.

Neurological Acceleration of stiff ODE solver with error monitoring and dynamic training			
Phase	**Standard Stiff ODE Solver Action**	**Neural Network Action**	**Description**
A	Always running	None	Collecting training data.
B	None	Initial training	Train *n* networks, one per constituent ODE.
C	Always running	Testing	Compare standard stiff ODE solver with neural network solution to train error networks, one per constituent ODE.
D	Always running	Update training	Update training of *n* ODE networks. Train *n* error networks giving neural network solution uncertainty.
E	Sometimes running	Operational with periodic training updates	**Operational.** When error networks indicate an error from the neural network solution that is less than the required threshold for each ODE use the neural network solution, otherwise use the regular stiff ODE solver. However, even when error network indicates an acceptable neural network solution periodically use the full stiff ODE solver solution to gather further training data to periodically update the training of all the networks.

Fig. 9. Strategy for applying a neural wrapper to accelerate the ODE solver.

Figure 9 shows the strategy that we used for applying a neural wrapper to accelerate the ODE solver. Figure 10 shows some example results for ozone after using a neural wrapper around an atmospheric chemistry ODE solver. The x-axis shows the actual ozone abundance as a volume mixing ratio (vmr) using the regular ODE solver without neural networks. The y-axis shows the ozone vmr inferred using the neural network solution. It can be seen that we have excellent agreement between the two solutions with a correlation coefficient of 1. The neural network has learned the behaviour of the ozone ODE very well. Without the adaptive error control the acceleration could be up to 200 times, with the full adaptive error control the acceleration was less, but usually at least a factor of two. Similarly, in Figure 11 the two panels below show the results for formaldehyde (HCHO) in the GMI model. The left panel shows the solution with SMVGear for level 1 at 01:00 UT and the right panel shows the corresponding solution using the neural network. As one would hope, the two results are almost indistinguishable.
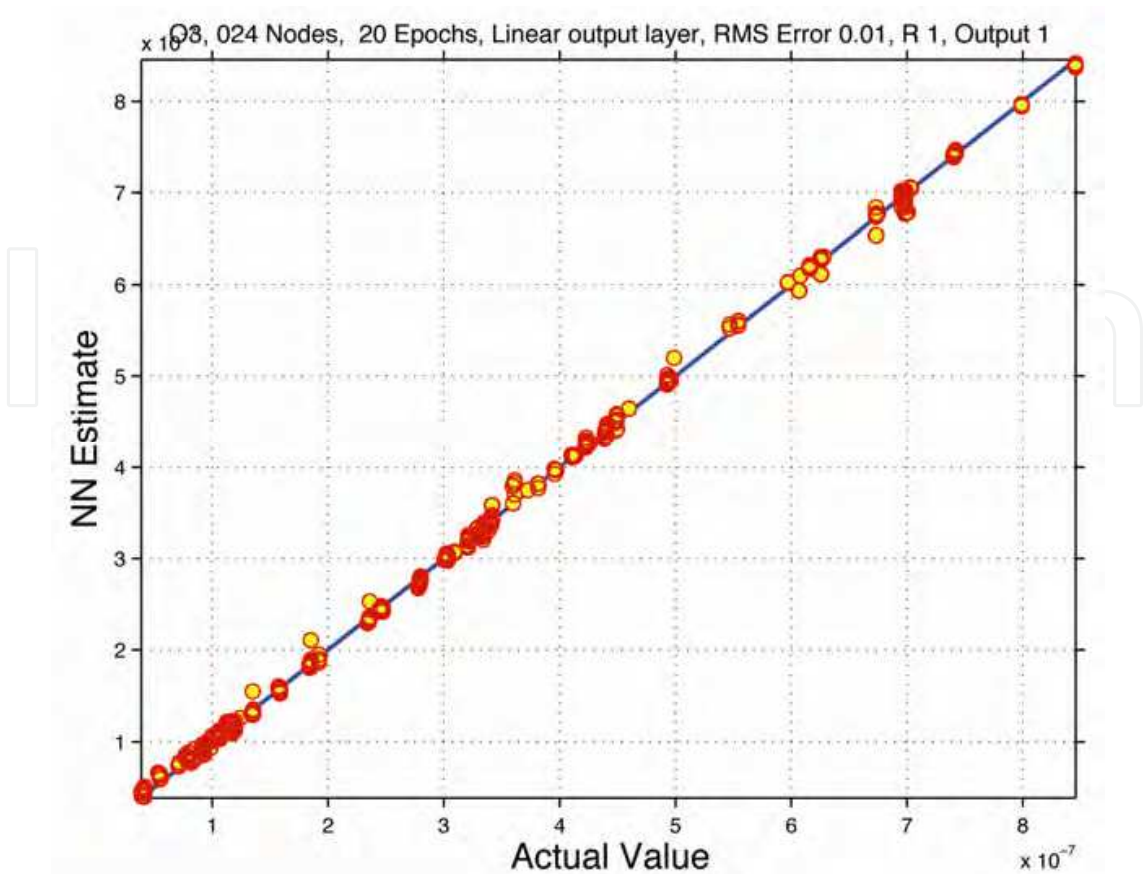
Fig. 10. Example results for using a neural wrapper around an atmospheric chemistry ODE solver. The x-axis shows the actual ozone v.m.r. using the regular ODE solver without neural networks. The y-axis shows the ozone v.m.r. inferred using the neural network solution. It can be seen that we have excellent agreement between the two solutions with a correlation coefficient of 1. The neural network has learned the behaviour of the ozone ODE very well.
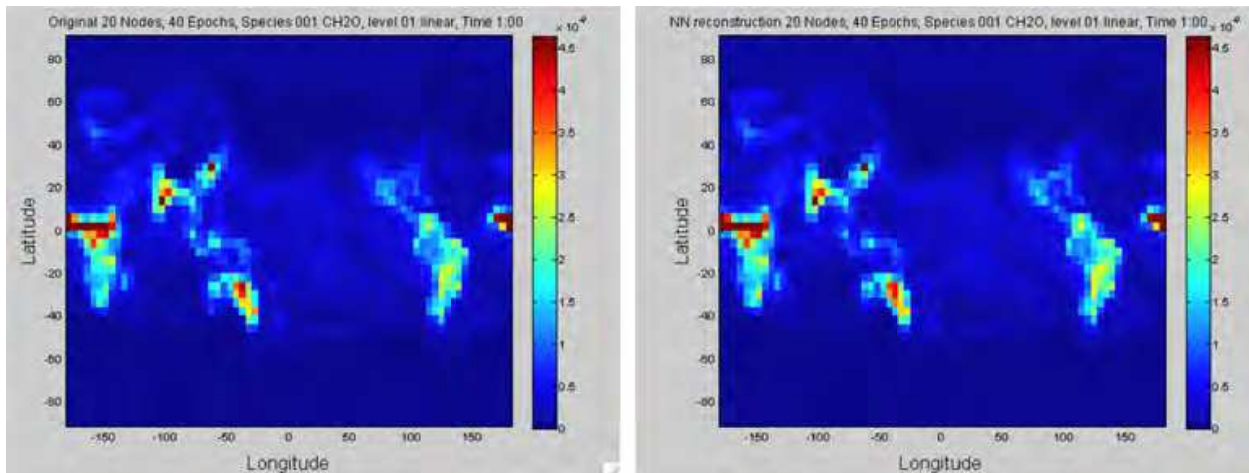

Fig. 11. The two panels below show the results for formaldehyde (HCHO) in the GMI model. The left panel shows the solution with SMVGear for level 1 at 01:00 UT and the right panel shows the corresponding solution using the neural network. As one would hope, the two results are almost indistinguishable.

4.6 Classification: Example from Detecting Drought Stress and Infection in Cacao

The source of chocolate, theobroma cacao (cacao), is an understory tropical tree (*Wood*, 2001). Cacao is intolerant to drought (*Belsky and Siebert*, 2003), and yields and production patterns are severely affected by periodic droughts and seasonal rainfall patterns. (*Bae et al.*, 2008) studied the molecular response of cacao to drought and have identified several genes responsive to drought stress (*Bailey et al.*, 2006). They have also been studying the response of cacao to colonization by an endophytic isolates of Trichoderma including Trichoderma hamatum, DIS 219b (*Bailey et al.*, 2006). One of the benefits to colonization Trichoderma hamatum isolate DIS 219b is tolerance to drought as mediated through plant growth promotion, specifically enhanced root growth (*Bae et al.*, 2008).

In characterizing the drought response of cacao considerable variation was observed in the response of individual seedlings depending upon the degree of drought stress applied (*Bae et al.*, 2008). In addition, although colonization by DIS 219b delayed the drought response, direct effects of DIS 219b on cacao gene expression in the absence of drought were difficult to identify (*Bae et al.*, 2008). The complexity of the DIS 219b/cacao plant microbe interaction overlaid on cacao's response to drought makes the system of looking at individual genes as a marker for either drought or endophyte inefficient.

There would be considerable utility in reliably predicting drought and endophyte stress from complex gene expression patterns, particularly as the endophyte lives within the plant without causing apparent phenotypic changes in the plant. Machine-learning models offer the possibility of highly accurate, automated predictions of plant stress from a variety of causes that may otherwise go undetected or be obscured by the complexity of plant responses to multiple environmental factors, to be considered status quo for plants in nature. We examined the ability of five different machine-learning approaches to predict drought stress and endophyte colonization in cacao: a naive Bayes classifier, decision trees (DTs), neural networks (NN), neuro-fuzzy inference (NFI), and support vector machine (SVM) classification. The results provided some support for the accuracy of machine-learning models in discerning endophyte colonization and drought stress. The best performance was by the neuro-fuzzy inference system and the support vector classifier that correctly identified 100% of the drought and endophyte stress samples. Of the two, the approaches the support vector classifier is likely to have the best generalization (wider applicability to data not previously seen in the training process).

Why did the SVM model outperform the four other machine learning approaches? We noted earlier that SVMs construct separating hyperplanes that maximize the margins between the different clusters in the training data set (the vectors that constrain the width of the margin are the support vectors). A good separation is achieved by those hyperplanes providing the largest distance between neighbouring classes, and in general, the larger the margin the better the generalization of the classifier.

When the points in neighbouring classes are separated by a nonlinear dividing line, rather than fitting nonlinear curves to the data, SVMs use a kernel function to map the data into a different space where a hyperplane can once more be used to do the separation. The kernel function may transform the data into a higher dimensional space to make it possible to perform the separation. The concept of a kernel mapping function is very powerful. It allows SVM models to perform separations even with very complex boundaries. Hence, we infer that, in the present application, the SVM model algorithmic process utilizes higher dimensional space to achieve superior predictive power.

For classification, the SVM algorithmic process offers an important advantage compared with neural network approaches. Specifically, neural networks can suffer from multiple local minima; in contrast, the solution to a support vector machine is global and unique. This characteristic may be partially attributed to the development process of these algorithms; SVMs were developed in the reverse order to the development of neural networks. SVMs evolved from the theory to implementation and experiments; neural networks followed a more heuristic path, from applications and extensive experimentation to theory.

In handling this data using traditional methods where individual gene responses are characterized as treatment effects, it was especially difficult to sort out direct effects of endophyte on gene expression over time or at specific time points. The differences between the responses of non-stressed plants with or without the endophyte were small and, after the zero time point, were highly variable. The general conclusion from this study was that colonization of cacao seedlings by the endophyte enhanced root growth resulting in increased drought tolerance but the direct effects of endophyte on cacao gene expression at the time points studied were minimal. Yet the neuro-fuzzy inference and support vector classification methods of analysis were able identify samples receiving these treatments correctly.

In this system, each gene in the plants genome is a potential sensor for the applied stress or treatment. It is not necessary that the genes response be significant in itself in determining the outcome of the plants response or that it be consistent in time or level of response. Since multiple genes are used in characterizing the response it is always the relative response in terms of the many other changes that are occurring at the same time as influenced by uncontrolled changes in the system that is important. With this study the treatments were controlled but variation in the genetic make up of each seedling (they were from segregating open pollinated seed) and minute differences in air currents within the chamber, soil composition, colonization levels, microbial populations within each pot and seedling, and even exact watering levels at each time point, all likely contributed to creating uncontrolled variation in the plants response to what is already a complex reaction to multiple factors (drought and endophyte). This type of variation makes accessing treatment responses using single gene approaches difficult and the prediction of cause due to effect in open systems almost impossible in complex systems.

5. Future Directions

We have seen the utility of machine learning for a suite of very diverse applications. These applications often help us make better use of existing data in a variety of ways. In parallel to the success of machine learning we also have the rapid development of publically available web services. So it is timely to combine both approached by providing online services that use machine learning for intelligent data fusion as part of a workflow that allows us to cross-calibrate multiple datasets. This obviously requires care to ensure the appropriate of datasets. However, if done carefully, this could greatly facilitate the production of seamless multi-year global records for a host of Earth science applications.

When it comes to dealing with inter-instrument biases in a consistent manner there is currently a gap in many space agencies' Earth science information systems. This could be addressed by providing an extensible and reusable open source infrastructure that gap that could be reused for multiple projects. A clear need for such an infrastructure would be for NASA's future Decadal Survey missions.

6. Summary

Machine learning has recently found many applications in the geosciences and remote sensing. These applications range from bias correction to retrieval algorithms, from code acceleration to detection of disease in crops. Machine-learning algorithms can act as "universal approximators", they can learn the behaviour of a system if they are given a comprehensive set of examples in a training dataset. Effective learning of the system's behaviour can be achieved even if it is multivariate and non-linear. An additional useful feature is that we do not need to know a priori the functional form of the system as required by traditional least-squares fitting, in other words they are non-parametric, non-linear and multivariate learning algorithms.

The uses of machine learning to date have fallen into three basic categories which are widely applicable across all of the Geosciences and remote sensing, the first two categories use machine learning for its regression capabilities, the third category uses machine learning for its classification capabilities. We can characterize the three application themes are as follows: First, where we have a theoretical description of the system in the form of a deterministic model, but the model is computationally *expensive*. In this situation, a machine-learning "wrapper" can be applied to the deterministic model providing us with a "code accelerator". Second, when we do not have a deterministic model but we have data available enabling us to empirically learn the behaviour of the system. Third, machine learning can be used for classification.

7. References

Anderson, J., Russell, J. M., Solomon, S. & Deaver, L. E. (2000) Halogen occultation experiment confirmation of stratospheric chlorine decreases in accordance with the montreal protocol. *Journal of Geophysical Research-Atmospheres,* 105**,** 4483-4490.

Atkinson, P. M. & Tatnall, A. R. L. (1997) Introduction: Neural networks in remote sensing. *International Journal of Remote Sensing,* 18**,** 699 - 709.

Bae, H., Kim, S. H., Kim, M. S., Sicher, R. C., Lary, D., Strem, M. D., Natarajan, S. & Bailey, B. A. (2008) The drought response of theobroma cacao (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses. *Plant Physiology and Biochemistry,* 46**,** 174-188.

Bailey, B. A., Bae, H., Strem, M. D., Roberts, D. P., Thomas, S. E., Crozier, J., Samuels, G. J., Choi, I. Y. & Holmes, K. A. (2006) Fungal and plant gene expression during the colonization of cacao seedlings by endophytic isolates of four trichoderma species. *Planta,* 224**,** 1449-1464.

Belsky, J. M. & Siebert, S. F. (2003) Cultivating cacao: Implications of sun-grown cacao on local food security and environmental sustainability. *Agriculture and Human Values,* 20**,** 277-285.

Bishop, C. M. (1995) *Neural networks for pattern recognition,* Oxford, Oxford University Press.

Bishop, C. M. (1998) *Neural networks and machine learning,* Berlin; New York, Springer.

Bonne, G. P., Stimpfle, R. M., Cohen, R. C., Voss, P. B., Perkins, K. K., Anderson, J. G., Salawitch, R. J., Elkins, J. W., Dutton, G. S., Jucks, K. W. & Toon, G. C. (2000) An examination of the inorganic chlorine budget in the lower stratosphere. *Journal of Geophysical Research-Atmospheres,* 105**,** 1957-1971.

Brown, M. E., Lary, D. J., Vrieling, A., Stathakis, D. & Mussa, H. (2008) Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS. *International Journal of Remote Sensing,* 29**,** 7141-7158.

Brown, M. E., Pinzon, J. E., Didan, K., Morisette, J. T. & Tucker, C. J. (2006) Evaluation of the consistency of long-term NDVI time series derived from AVHRR, spot-vegetation, seawifs, MODIS and landsat etm+. *IEEE Transactions Geoscience and Remote Sensing,* 44**,** 1787-1793.

Carpenter, G. A., Gjaja, M. N., Gopal, S. & Woodcock, C. E. (1997) Art neural networks for remote sensing: Vegetation classification from landsat tm and terrain data. *IEEE Transactions on Geoscience and Remote Sensing,* 35**,** 308-325.

Caselli, M., Trizio, L., De Gennaro, G. & Ielpo, P. (2009) A simple feedforward neural network for the pm10 forecasting: Comparison with a radial basis function network and a multivariate linear regression model. *Water Air and Soil Pollution,* 201**,** 365-377.

Chen, P. H., Fan, R. E. & Lin, C. J. (2006) A study on smo-type decomposition methods for support vector machines. *Ieee Transactions on Neural Networks,* 17**,** 893-908.

Chevallier, F., Cheruy, F., Scott, N. A. & Chedin, A. (1998) A neural network approach for a fast and accurate computation of a longwave radiative budget. *Journal of Applied Meteorology,* 37**,** 1385-1397.

Comrie, A. C. (1997) Comparing neural networks and regression models for ozone forecasting. *Journal of the Air & Waste Management Association,* 47**,** 653-663.

Dufour, G., Nassar, R., Boone, C. D., Skelton, R., Walker, K. A., Bernath, P. F., Rinsland, C. P., Semeniuk, K., Jin, J. J., Mcconnell, J. C. & Manney, G. L. (2006) Partitioning between the inorganic chlorine reservoirs HCl and ClONO2 during the arctic winter 2005 from the ace-fts. *Atmospheric Chemistry and Physics,* 6**,** 2355-2366.

Eklundh, L. & Olsson, L. (2003) Vegetation index trends for the african sahel 1982-1999. *Geophysical Research Letters,* 30.

Eyring, V., Butchart, N., Waugh, D. W., Akiyoshi, H., Austin, J., Bekki, S., Bodeker, G. E., Boville, B. A., Bruhl, C., Chipperfield, M. P., Cordero, E., Dameris, M., Deushi, M., Fioletov, V. E., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Grewe, V., Jourdain, L., Kinnison, D. E., Mancini, E., Manzini, E., Marchand, M., Marsh, D. R., Nagashima, T., Newman, P. A., Nielsen, J. E., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Schraner, M., Shepherd, T. G., Shibata, K., Stolarski, R. S., Struthers, H., Tian, W. & Yoshiki, M. (2006) Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past. *Journal of Geophysical Research-Atmospheres,* 111.

Eyring, V., Waugh, D. W., Bodeker, G. E., Cordero, E., Akiyoshi, H., Austin, J., Beagley, S. R., Boville, B. A., Braesicke, P., Bruhl, C., Butchart, N., Chipperfield, M. P., Dameris, M., Deckert, R., Deushi, M., Frith, S. M., Garcia, R. R., Gettelman, A., Giorgetta, M. A., Kinnison, D. E., Mancini, E., Manzini, E., Marsh, D. R., Matthes, S., Nagashima, T., Newman, P. A., Nielsen, J. E., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Schraner, M., Scinocca, J. F., Semeniuk, K., Shepherd, T. G., Shibata, K., Steil, B., Stolarski, R. S., Tian, W. & Yoshiki, M. (2007) Multimodel projections of stratospheric ozone in the 21st century. *Journal of Geophysical Research-Atmospheres,* 112.

Fan, R. E., Chen, P. H. & Lin, C. J. (2005) Working set selection using second order information for training support vector machines. *Journal of Machine Learning Research,* 6**,** 1889-1918.

Froidevaux, L., Jiang, Y. B., Lambert, A., Livesey, N. J., Read, W. G., Waters, J. W., Fuller, R. A., Marcy, T. P., Popp, P. J., Gao, R. S., Fahey, D. W., Jucks, K. W., Stachnik, R. A., Toon, G. C., Christensen, L. E., Webster, C. R., Bernath, P. F., Boone, C. D., Walker, K. A., Pumphrey, H. C., Harwood, R. S., Manney, G. L., Schwartz, M. J., Daffer, W. H., Drouin, B. J., Cofield, R. E., Cuddy, D. T., Jarnot, R. F., Knosp, B. W., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P. & Wagner, P. A. (2008) Validation of Aura microwave limb sounder HCl measurements. *Journal of Geophysical Research-Atmospheres,* 113.

Froidevaux, L., Livesey, N. J., Read, W. G., Jiang, Y. B. B., Jimenez, C., Filipiak, M. J., Schwartz, M. J., Santee, M. L., Pumphrey, H. C., Jiang, J. H., Wu, D. L., Manney, G. L., Drouin, B. J., Waters, J. W., Fetzer, E. J., Bernath, P. F., Boone, C. D., Walker, K. A., Jucks, K. W., Toon, G. C., Margitan, J. J., Sen, B., Webster, C. R., Christensen, L. E., Elkins, J. W., Atlas, E., Lueb, R. A. & Hendershot, R. (2006a) Early validation analyses of atmospheric profiles from eos MLS on the Aura satellite. *IEEE Transactions on Geoscience and Remote Sensing,* 44**,** 1106-1121.

Froidevaux, L., Livesey, N. J., Read, W. G., Salawitch, R. J., Waters, J. W., Drouin, B., Mackenzie, I. A., Pumphrey, H. C., Bernath, P., Boone, C., Nassar, R., Montzka, S., Elkins, J., Cunnold, D. & Waugh, D. (2006b) Temporal decrease in upper atmospheric chlorine. *Geophysical Research Letters,* 33.

Gallo, K. P., Ji, L., Reed, B. C., Dwyer, J. & Eidenshink, J. C. (2003) Comparison of MODIS and AVHRR 16-day normalized difference vegetation index composite data. *Geophysical Research Letters,* 31**,** L07502-5.

Gardner, M. W. & Dorling, S. R. (1999) Neural network modelling and prediction of hourly NOx and NO2 concentrations in urban air in london. *Atmospheric Environment,* 33**,** 709-719.

Gunson, M. R., Abrams, M. C., Lowes, L. L., Mahieu, E., Zander, R., Rinsland, C. P., Ko, M. K. W., Sze, N. D. & Weisenstein, D. K. (1994) Increase in levels of stratospheric chlorine and fluorine loading between 1985 and 1992. *Geophysical Research Letters,* 21**,** 2223-2226.

Haykin, S. (2001a) *Kalman filtering and neural networks*, Wiley-Interscience.

Haykin, S. S. (1994) *Neural networks : A comprehensive foundation,* New York, Toronto, Macmillan.

Haykin, S. S. (2001b) *Kalman filtering and neural networks,* New York, Wiley.

Haykin, S. S. (2007) *New directions in statistical signal processing : From systems to brain,* Cambridge, Mass., MIT Press.

Hyyppa, J., Hyyppa, H., Inkinen, M., Engdahl, M., Linko, S. & Zhu, Y. H. (1998) Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management.* Lake Buena Vista, Florida.

Lary, D. J. & Aulov, O. (2008) Space-based measurements of HCl: Intercomparison and historical context. *Journal of Geophysical Research-Atmospheres,* 113.

Lary, D. J., Muller, M. D. & Mussa, H. Y. (2004) Using neural networks to describe tracer correlations. *Atmospheric Chemistry and Physics,* 4**,** 143-146.

Lary, D. J., Remer, L., Paradise, S., Macneill, D. & Roscoe, B. (2009) Machine learning and bias correction of MODIS aerosol optical depth. *IEEE Trans. on Geoscience and Remote Sensing*

Lary, D. J., Waugh, D. W., Douglass, A. R., Stolarski, R. S., Newman, P. A. & Mussa, H. (2007) Variations in stratospheric inorganic chlorine between 1991 and 2006. *Geophysical Research Letters,* 34.

Levenberg, K. (1944) A method for the solution of certain problems in least squares. *Quart. Appl. Math.,* 2**,** 164-168.

Los, S. O. (1998) Estimation of the ratio of sensor degradation between noaa AVHRR channels 1 and 2 from monthly NDVI composites. *IEEE Transactions on Geoscience and Remote Sensing,* 36**,** 206-213.

Marquardt, D. W. (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics,* 11**,** 431-441.

Marquardt, D. W. (1979) Citation classic - algorithm for least-squares estimation of non-linear parameters. *Current Contents/Engineering Technology & Applied Sciences***,** 14-14.

Mchugh, M., Magill, B., Walker, K. A., Boone, C. D., Bernath, P. F. & Russell, J. M. (2005) Comparison of atmospheric retrievals from ace and HALOE. *Geophysical Research Letters,* 32.

Michelsen, H. A., Salawitch, R. J., Gunson, M. R., Aellig, C., Kampfer, N., Abbas, M. M., Abrams, M. C., Brown, T. L., Chang, A. Y., Goldman, A., Irion, F. W., Newchurch, M. J., Rinsland, C. P., Stiller, G. P. & Zander, R. (1996) Stratospheric chlorine partitioning: Constraints from shuttle-borne measurements of [HCl], [clno3], and [ClO]. *Geophysical Research Letters,* 23**,** 2361-2364.

Miura, T., Huete, A. & Yoshioka, H. (2006) An empirical investigation of cross-sensor relationships of NDVI and red/near-infrared reflectance using eo-1 hyperion data. *Remote Sensing of Environment,* 100**,** 223-236.

Moré, J. J. (1977) The levenberg-marquardt algorithm: Implementation and theory. IN Watson, G. A. (Ed.) *Numerical analysis.* Springer Verlag.

Nassar, R., Bernath, P. F., Boone, C. D., Clerbaux, C., Coheur, P. F., Dufour, G., Froidevaux, L., Mahieu, E., Mcconnell, J. C., Mcleod, S. D., Murtagh, D. P., Rinsland, C. P., Semeniuk, K., Skelton, R., Walker, K. A. & Zander, R. (2006) A global inventory of stratospheric chlorine in 2004. *Journal of Geophysical Research-Atmospheres,* 111.

Newman, P. A., Nash, E. R., Kawa, S. R., Montzka, S. A. & Schauffler, S. M. (2006) When will the antarctic ozone hole recover? *Geophysical Research Letters,* 33.

Rinsland, C. P., Gunson, M. R., Salawitch, R. J., Michelsen, H. A., Zander, R., Newchurch, M. J., Abbas, M. M., Abrams, M. C., Manney, G. L., Chang, A. Y., Irion, F. W., Goldman, A. & Mahieu, E. (1996) ATMOS/atlas-3 measurements of stratospheric chlorine and reactive nitrogen partitioning inside and outside the november 1994 antarctic vortex. *Geophysical Research Letters,* 23**,** 2365-2368.

Russell, J. M., Deaver, L. E., Luo, M. Z., Park, J. H., Gordley, L. L., Tuck, A. F., Toon, G. C., Gunson, M. R., Traub, W. A., Johnson, D. G., Jucks, K. W., Murcray, D. G., Zander, R., Nolt, I. G. & Webster, C. R. (1996) Validation of hydrogen chloride measurements made by the halogen occultation experiment from the UARS platform. *Journal of Geophysical Research-Atmospheres,* 101**,** 10151-10162.

Santee, M. L., Mackenzie, I. A., Manney, G. L., Chipperfield, M. P., Bernath, P. F., Walker, K. A., Boone, C. D., Froidevaux, L., Livesey, N. J. & Waters, J. W. (2008) A study of stratospheric chlorine partitioning based on new satellite measurements and modeling. *Journal of Geophysical Research-Atmospheres,* 113.

Scholkopf, B., Smola, A. J., Williamson, R. C. & Bartlett, P. L. (2000) New support vector algorithms. *Neural Computation,* 12**,** 1207-1245.

Sen, B., Osterman, G. B., Salawitch, R. J., Toon, G. C., Margitan, J. J., Blavier, J. F., Chang, A. Y., May, R. D., Webster, C. R., Stimpfle, R. M., Bonne, G. P., Voss, P. B., Perkins, K. K., Anderson, J. G., Cohen, R. C., Elkins, J. W., Dutton, G. S., Hurst, D. F., Romashkin, P. A., Atlas, E. L., Schauffler, S. M. & Loewenstein, M. (1999) The budget and partitioning of stratospheric chlorine during the 1997 arctic summer. *Journal of Geophysical Research-Atmospheres,* 104**,** 26653-26665.

Slayback, D. A., Pinzon, J. E., Los, S. O. & Tucker, C. J. (2003) Northern hemisphere photosynthetic trends 1982-99. *Global Change Biology,* 9**,** 1-15.

Smola, A. J. & Scholkopf, B. (2004) A tutorial on support vector regression. *Statistics and Computing,* 14**,** 199-222.

Solomon, S., Intergovernmental Panel on Climate Change. & Intergovernmental Panel on Climate Change. Working Group I. (2007) *Climate change 2007 : The physical science basis : Contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change,* Cambridge ; New York, Cambridge University Press.

Steven, M. D., Malthus, T. J., Baret, F., Xu, H. & Chopping, M. J. (2003) Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment,* 88**,** 412-422.

Teillet, M., Staenz, K. & Williams, D. J. (1997) Effects of spectral, spatial and radiometric characteristics on remote sensing vegetation indices of forested regions. *Remote Sensing of Environment,* 61**,** 139-149.

Trishchenko, A. P., Cihlar, J. & Li, Z. (2002) Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sensing of Environment,* 81**,** 1-18.

Tucker, C. J., Slayback, D. A., Pinzon, J. E., Los, S. O., Myneni, R. B. & Taylor, M. G. (2001) Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. *International Journal of Biometeorology,* 45**,** 184-190.

Van Leeuwen, W., Orr, B. J., Marsh, S. E. & Herrmann, S. M. (2006) Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. *Remote Sensing of Environment,* 100**,** 67-81.

Vapnik, V. N. (1995) *The nature of statistical learning theory,* New York, Springer.

Vapnik, V. N. (1998) *Statistical learning theory,* New York, Wiley.

Vapnik, V. N. (2000) *The nature of statistical learning theory,* New York, Springer.

Voss, P. B., Stimpfle, R. M., Cohen, R. C., Hanisco, T. F., Bonne, G. P., Perkins, K. K., Lanzendorf, E. J., Anderson, J. G., Salawitch, R. J., Webster, C. R., Scott, D. C., May, R. D., Wennberg, P. O., Newman, P. A., Lait, L. R., Elkins, J. W. & Bui, T. P. (2001) Inorganic chlorine partitioning in the summer lower stratosphere: Modeled and measured [ClONO2]/[HCl] during polaris. *Journal of Geophysical Research-Atmospheres,* 106**,** 1713-1732.

Waugh, D. W. & Eyring, V. (2008) Quantitative performance metrics for stratospheric-resolving chemistry-climate models. *Atmospheric Chemistry and Physics,* 8**,** 5699-5713.

Webster, C. R., May, R. D., Jaegle, L., Hu, H., Sander, S. P., Gunson, M. R., Toon, G. C., Russell, J. M., Stimpfle, R. M., Koplow, J. P., Salawitch, R. J. & Michelsen, H. A. (1994) Hydrochloric-acid and the chlorine budget of the lower stratosphere. *Geophysical Research Letters,* 21**,** 2575-2578.

Wmo (2006) Scientific assessment of ozone depletion: 2006. WMO Global Ozone Res. and Monitor. Proj., Geneva.

Wood, G. A. R. (2001) *CACAO,* 620.

Yi, J. S. & Prybutok, V. R. (1996) A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area. *Environmental Pollution,* 92**,** 349-357.

Zander, R., Gunson, M. R., Farmer, C. B., Rinsland, C. P., Irion, F. W. & Mahieu, E. (1992) The 1985 chlorine and fluorine inventories in the stratosphere based on ATMOS observations at 30-degrees north latitude. *Journal of Atmospheric Chemistry,* 15**,** 171-186.

Zander, R., Mahieu, E., Gunson, M. R., Abrams, M. C., Chang, A. Y., Abbas, M., Aellig, C., Engel, A., Goldman, A., Irion, F. W., Kampfer, N., Michelsen, H. A., Newchurch, M. J., Rinsland, C. P., Salawitch, R. J., Stiller, G. P. & Toon, G. C. (1996) The 1994 northern midlatitude budget of stratospheric chlorine derived from ATMOS/atlas-3 observations. *Geophysical Research Letters,* 23**,** 2357-2360.

Geoscience and Remote Sensing New Achievements

Edited by Pasquale Imperatore and Daniele Riccio

Our planet is nowadays continuously monitored by powerful remote sensors operating in wide portions of the electromagnetic spectrum. Our capability of acquiring detailed information on the environment has been revolutionized by revealing its inner structure, morphology and dynamical changes. The way we now observe and study the evolution of the Earth's status has even radically influenced our perception and conception of the world we live in. The aim of this book is to bring together contributions from experts to present new research results and prospects of the future developments in the area of geosciences and remote sensing; emerging research directions are discussed. The volume consists of twenty-six chapters, encompassing both theoretical aspects and application-oriented studies. An unfolding perspective on various current trends in this extremely rich area is offered. The book chapters can be categorized along different perspectives, among others, use of active or passive sensors, employed technologies and configurations, considered scenario on the Earth, scientific research area involved in the studies.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds