



Database to determine global changes in vertically resolved stratospheric O₃

Greg Bodeker

Bodeker Scientific, Alexandra, New Zealand

Birgit Hassler

CIRES/NOAA, Boulder, USA

**SPARC/IOC/WMO-IGAO workshop on
Past Changes in the Vertical Distribution of Ozone**

Outline

- Description of the input data sets.
- 5 different coordinate systems for the data.
- Homogenizing the input data.
- Creating a global gap-free database suitable for long-term trend detection.

Source data for the database

Version 1.1.0.6 of the BDBP

Available from:

www.bodekerscientific.com/Data/the-bdbp

Includes as input:

- LIMS - 1978 to 1979
- SAGE I - 1979 to 1981
- SAGE II - 1984 to 2005
- ILAS - 1996 to 1997
- AURA MLS - 2004 to 2009 (still in progress)
- All ozonesondes from WOUDC, NDACC and elsewhere.
- POAM II - 1993 to 1996
- POAM III - 1998 to 2005
- HALOE - 1991 to 2005
- ILAS II - 2003

The database includes more than just ozone but for the purposes of this talk we concentrate on ozone only.

Source data for the database

Plans for the future

Next we intend to add data from:

- GOMOS
- MIPAS
- SCIAMACHY
- OSIRIS
- ACE-FTS

Data Grids

The data are stored in 5 different grid formats. Data are interpolated from the source data files onto the grid levels. Each source datum is added to all 5 grids.

Latitude/altitude: 70 altitudes extending from 1 to 70 km.

Latitude/pressure: 70 pressure levels $p = p_0 e^{(-i/7)}$ $i=1$ to 70.

Equivalent latitude/theta: Levels 300, 315, 330, 350, 400, 450, 550 and 650K.

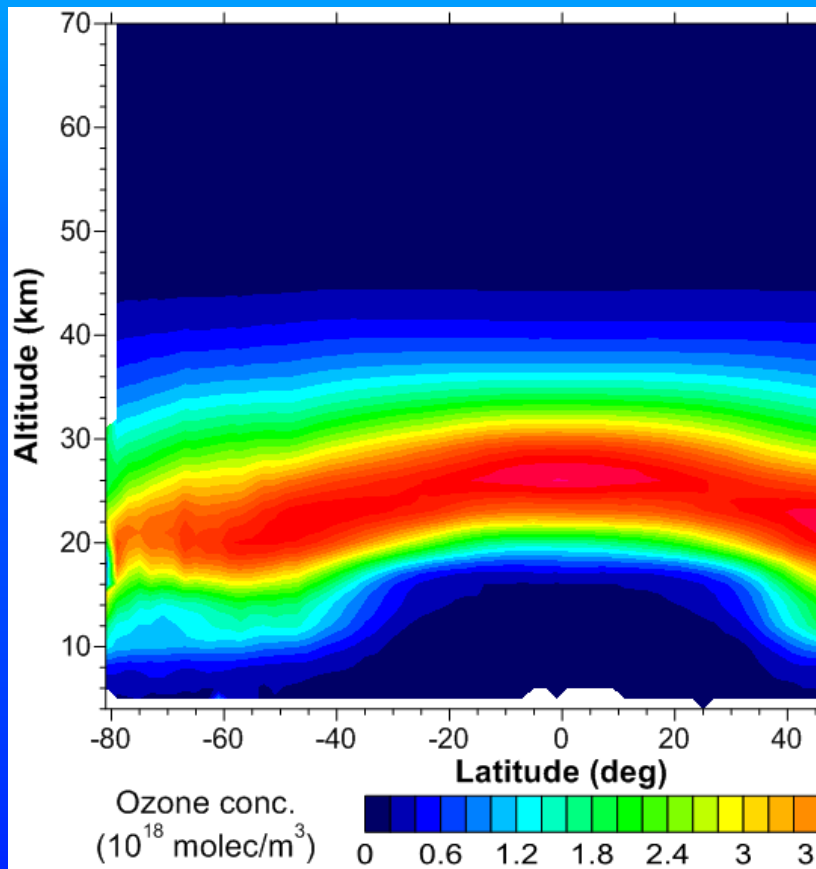
Latitude/altitude w.r.t. tropopause altitude: 70 altitudes above/below the tropopause with higher density close to the tropopause.

Latitude/pressure w.r.t. tropopause pressure: 70 pressure levels above/below the tropopause with higher density close to the tropopause.

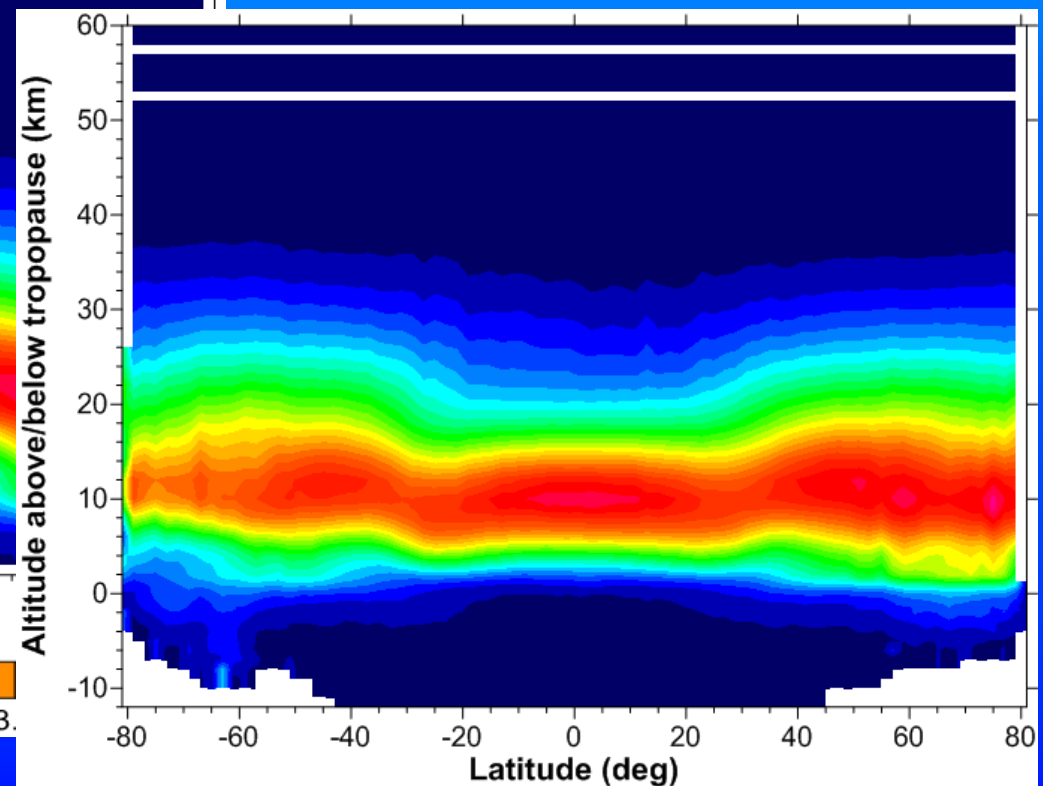
Example climatologies

HALOE annual mean climatologies

Standard altitude



Altitude with respect to local tropopause height



Homogenizing the input data

We begin by assuming that the SAGE II data require no correction.

Drifts or discontinuities in the ozonesonde data are identified by changes in their normalization factors (independent TCO/ozonesonde TCO) which are stored along with each ozonesonde measurement. We apply these normalization factors to the ozonesonde data to create homogeneous time series.

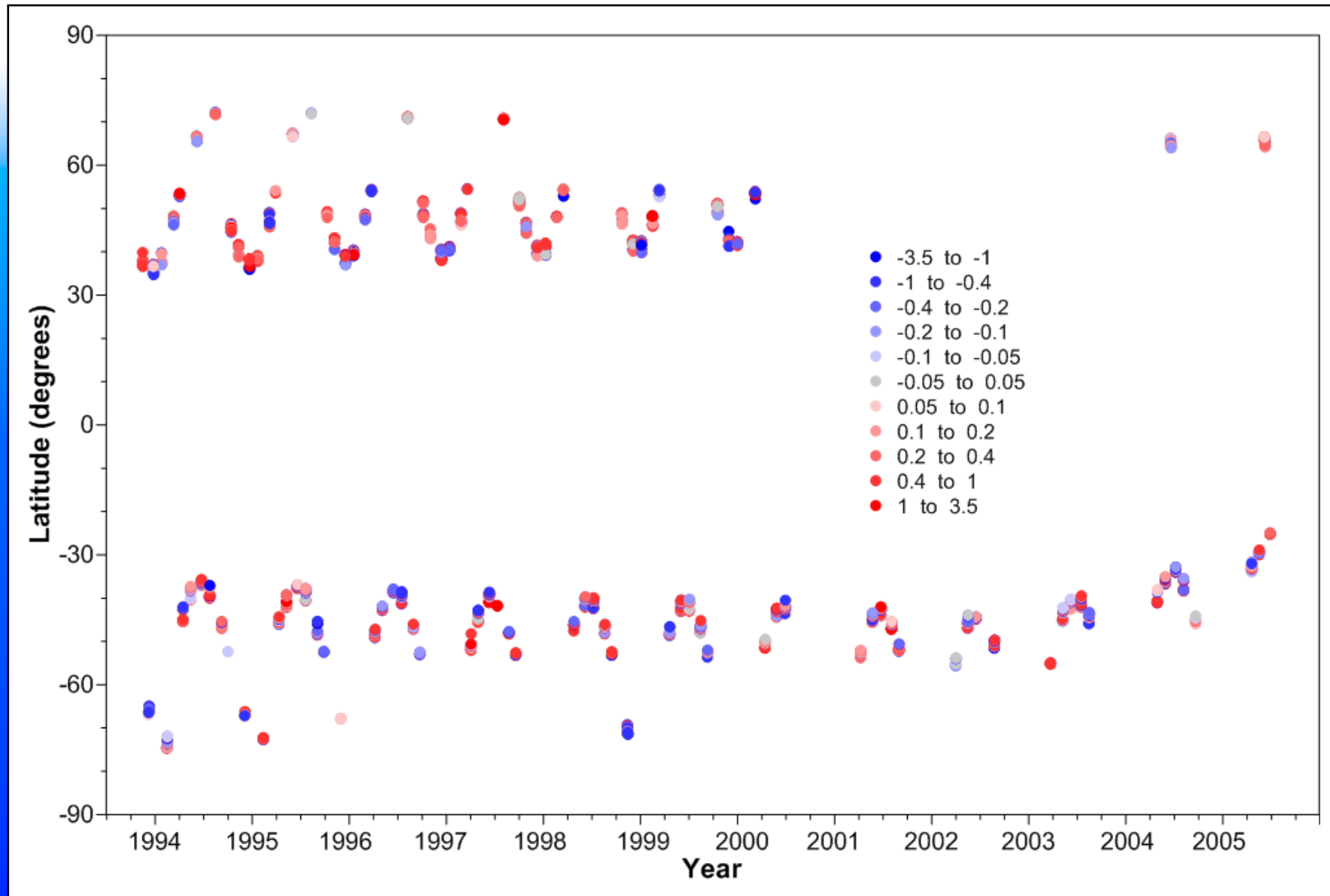
Homogenizing the input data

Corrected ozonesonde data are merged with the SAGE II data to create a 'gold standard'. On each altitude, pressure or potential temperature surface, we then scan the remaining data sets to find which gives the greatest number of coincidences within prescribed latitude, longitude and time ranges (typically 3° latitude, 20° longitude, 6 hours time). Difference pairs are then calculated. To the latitude/time field of difference pairs on a given level we fit:

$$\text{Diff} = A + B^*t$$

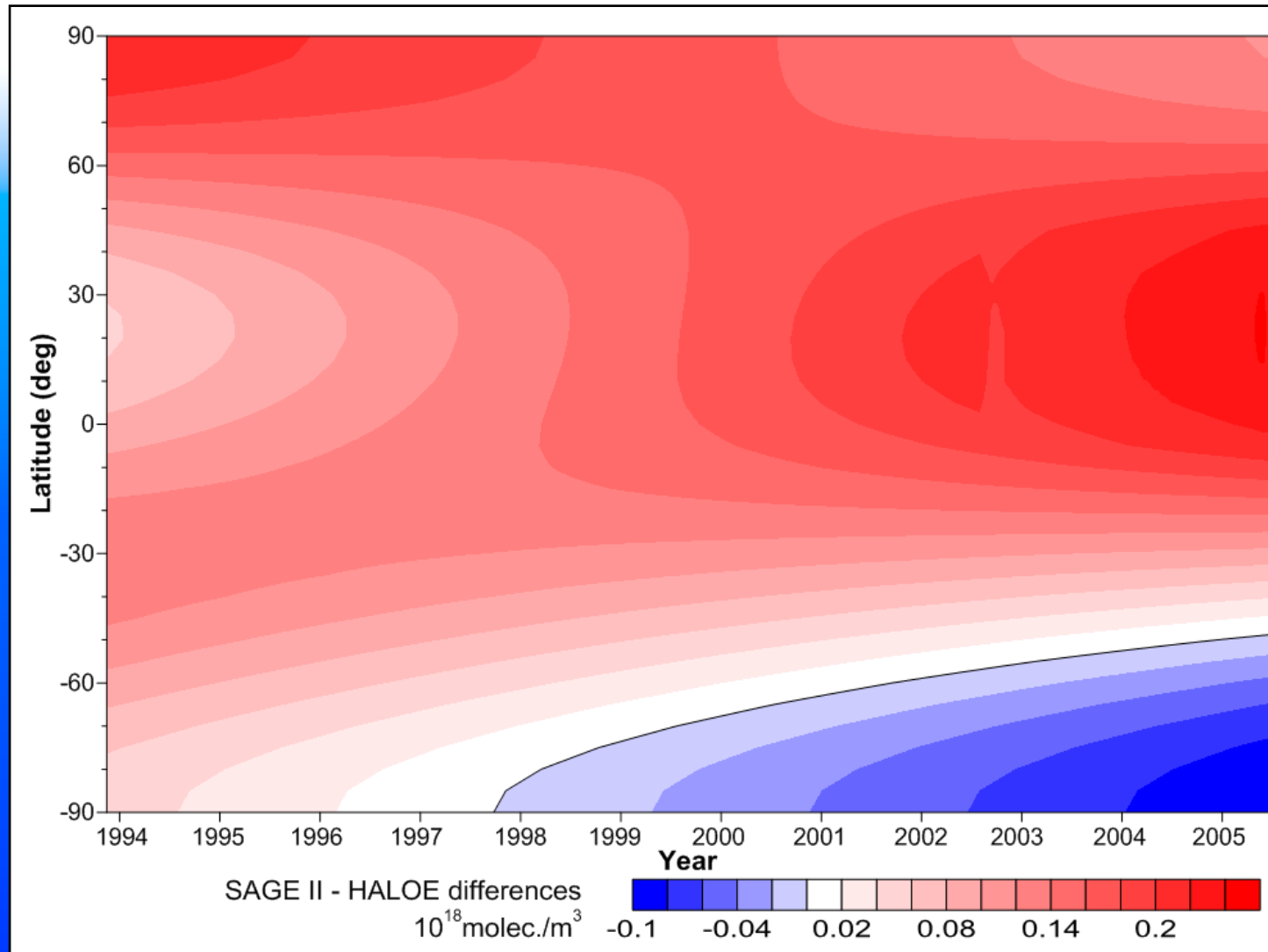
Where **A** and **B** are expanded in Fourier terms to account for seasonality in the differences and in Legendre polynomials to account for latitudinal structure in the differences...

Example SAGE II and HALOE coincidences



Number density differences at 20 km in 10^{18} molec./m³

Example SAGE II and HALOE coincidences



This filled field can now be used to correct HALOE globally.
This is done independently at each level.

Homogenizing the input data

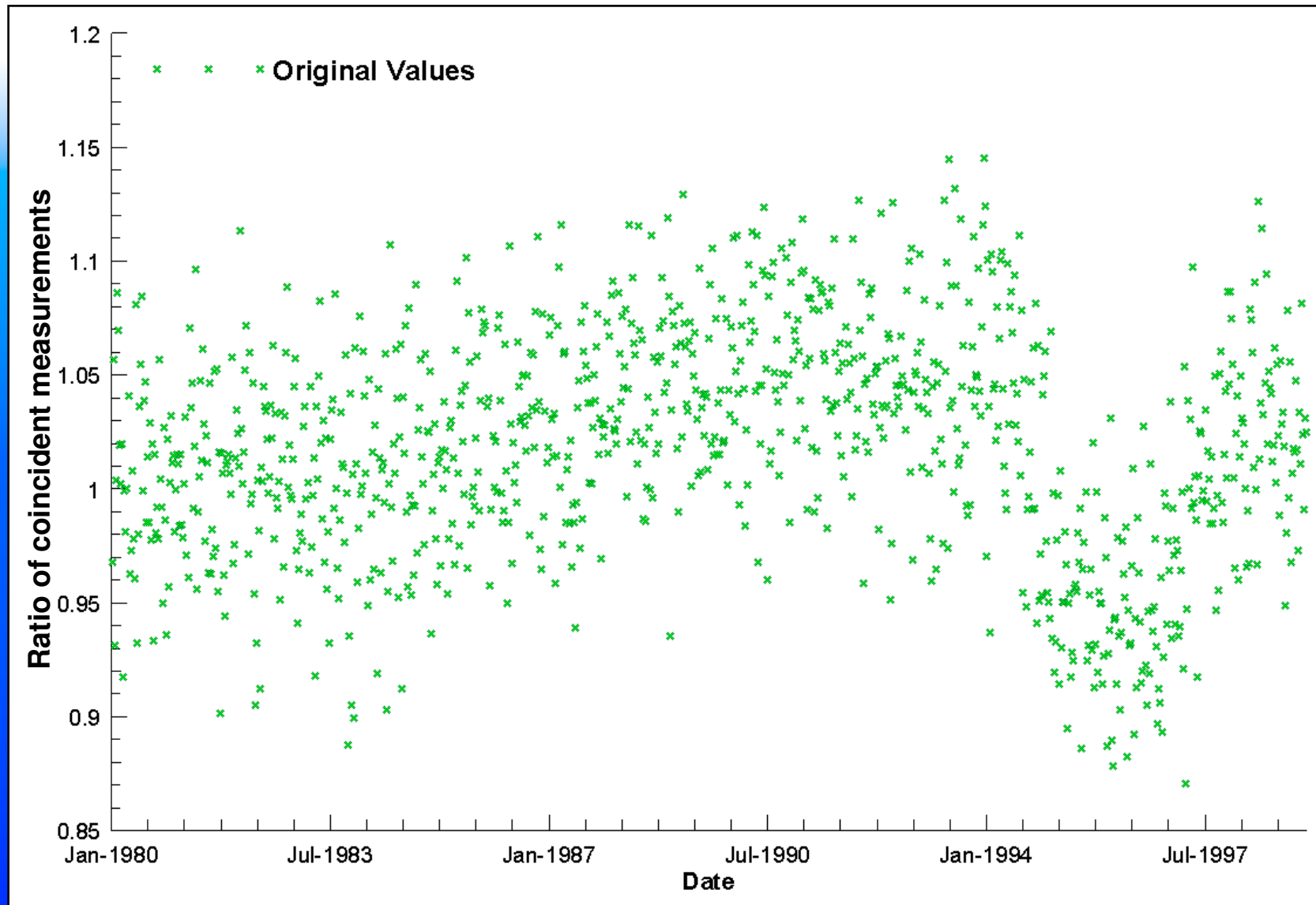
The corrected HALOE data get added to SAGE II + ozonesondes to make a new gold standard. This is repeated until all data sets have been ingested into the gold standard.

We have one additional weapon in our arsenal which we intend to integrate with the regression fit to differences method just described. An algorithm based on:

Alexandersson, H. and Moberg, A., Homogenization of Swedish temperature data. Part I: homogeneity test for linear trends, International Journal of Climatology, 17, 25-34, 1997.

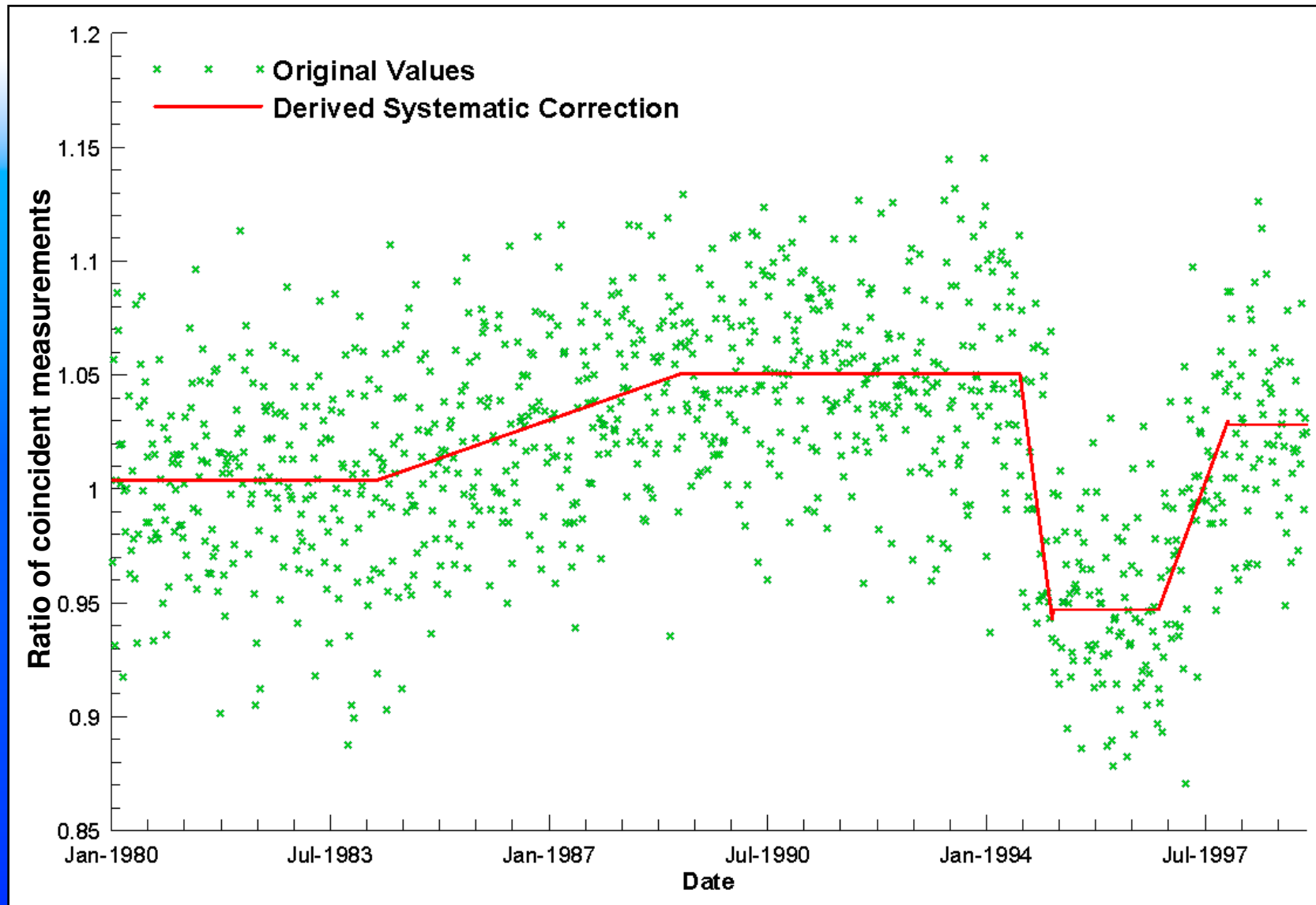
has been developed to detect and remove offsets and drifts in difference or ratio time series. This would extend the regression model approach to account for systematic discontinuities in differences or ratios between data time series.

Example of correcting ratio time series



Note: synthetic data - just an example

Example of correcting ratio time series



Red line gives *systematic* structure.

Homogenizing the input data

The structure captured by the red line would replace the trend and offset basis functions in our regression model fit to the differences. We just need to figure out how to expand that method to account for seasonality and dependence on latitude e.g. using Fourier and Legendre expansions.

These homogenization techniques are still being developed and have not yet been implemented in what is shown below. For now, we have a much simpler system where, when calculating monthly means from the database, we iteratively discard values that are more than 3σ away from the mean.

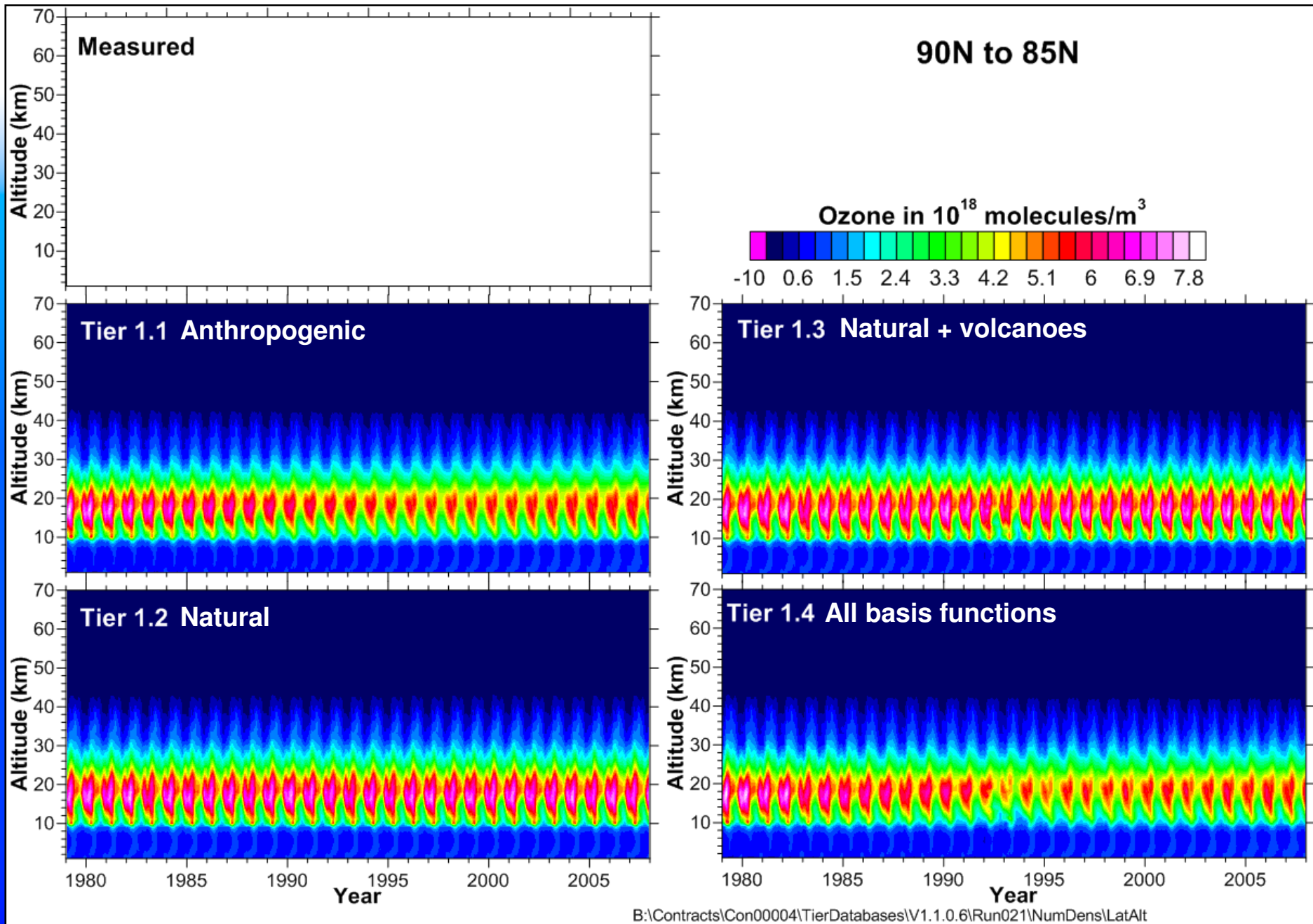
Creating a globally filled data set

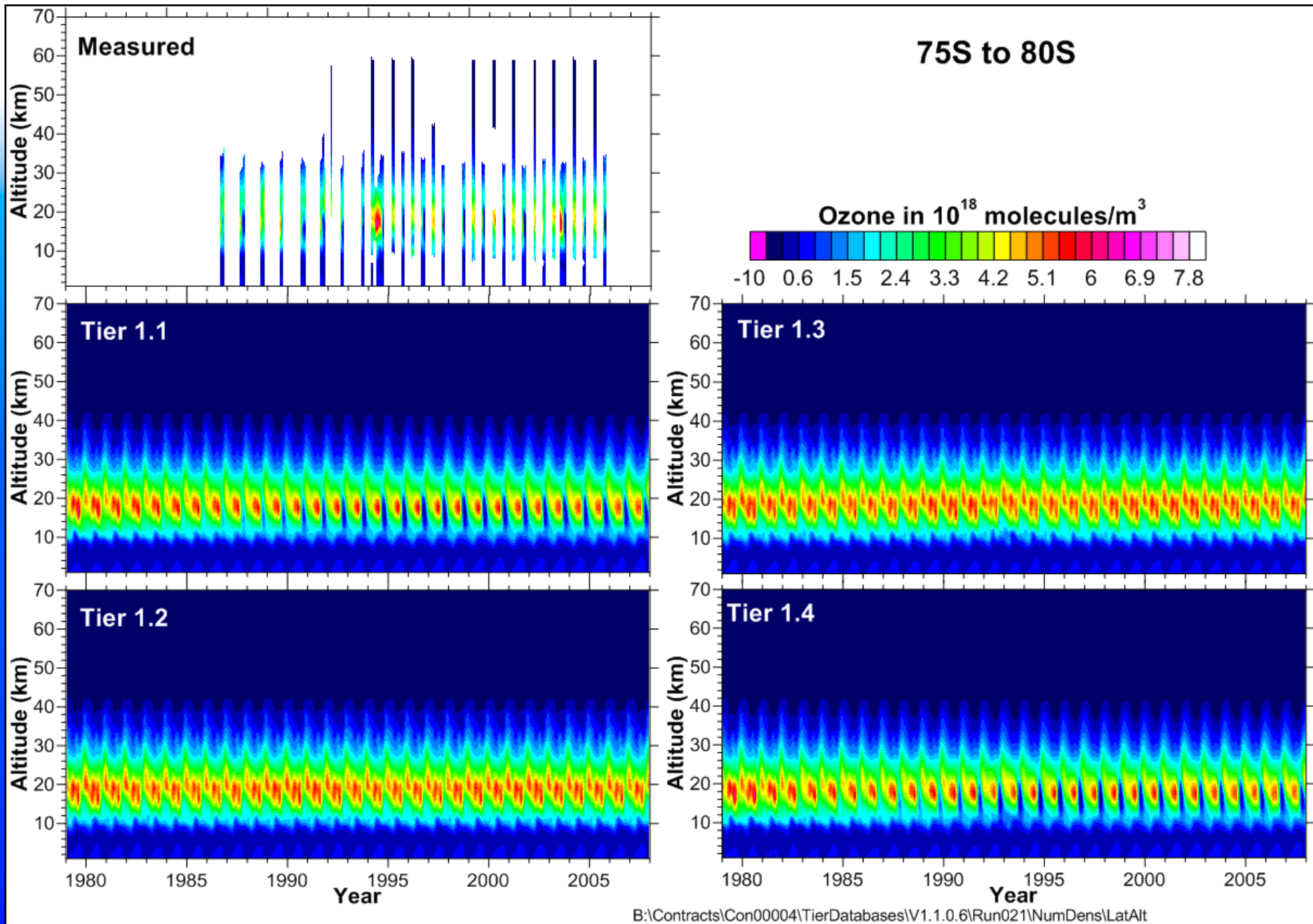
Least squares regression model is fitted to the monthly means:

Ozone(t,l) =

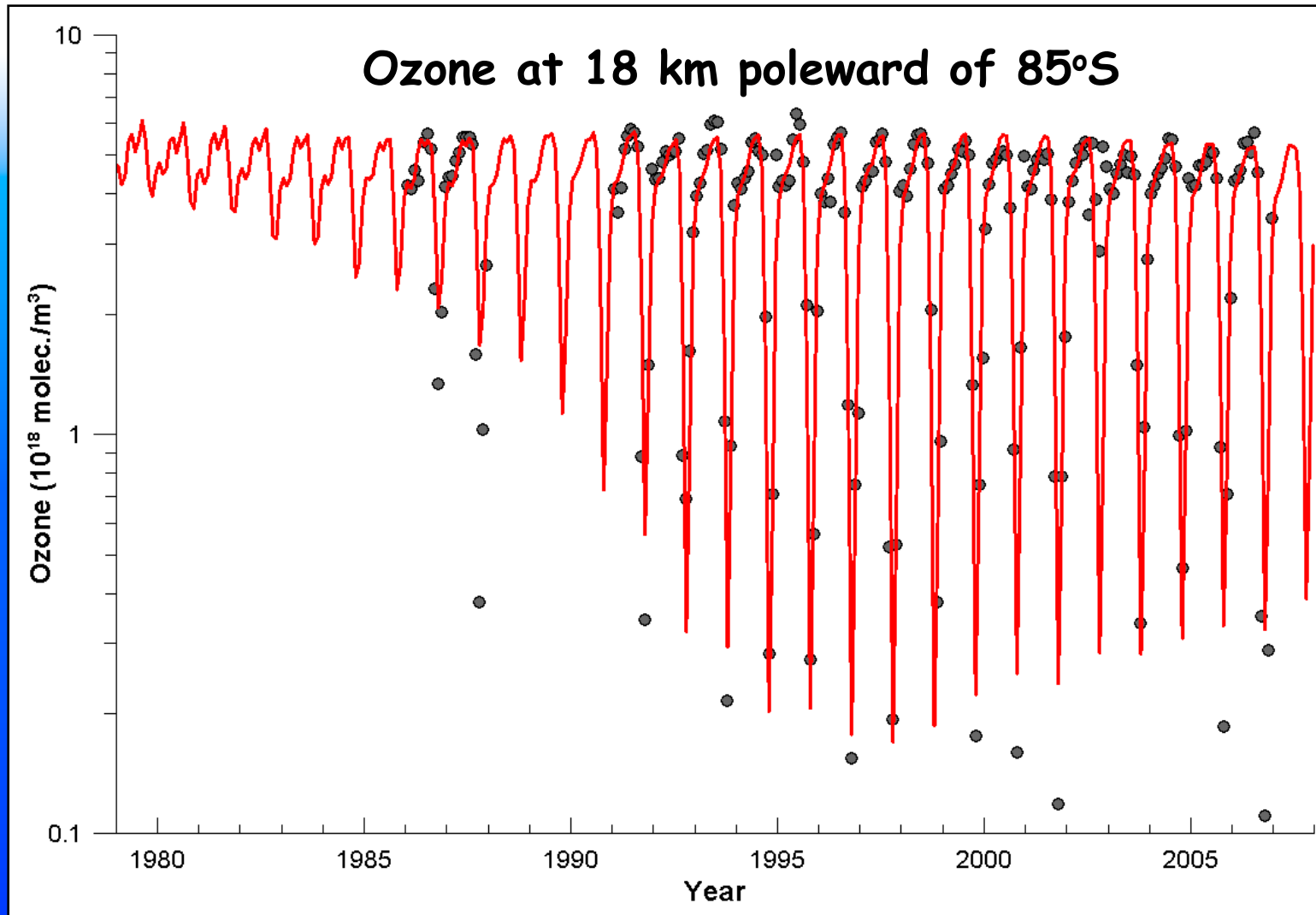
$A(t,l) +$	Offset and seasonal cycle
$B(t,l) \times t +$	Linear trend
$C(t,l) \times \text{EESC}(t,\Gamma) +$	Age of air dependent EESC (stratosphere only)
$D(t,l) \times \text{QBO}(t)$	Quasi-biennial Oscillation
$E(t,l) \times \text{QBO}_{\text{orth}}(t) +$	Orthogonalized QBO
$F(t,l) \times \text{ENSO}(t) +$	El Niño-Southern Oscillation
$G(t,l) \times \text{Solar}(t) +$	Solar cycle
$H(t,l) \times \text{Pinatubo}(t) +$	Mt. Pinatubo eruption
$R(t)$	

Fit coefficients are expanded in Fourier series to account for seasonality (t) and in Legendre polynomials for dependence on latitude (l). So a single fit is done (typically 180 fit coeff.) to the data at each level. Some examples of the results...





Capturing the deep ozone depletion over Antarctica



When fitting regression model, transform O_3 to log scale when $O_3 < 1.0$ and inverse transform back.

Conclusions

- We have a database (BDBP) that includes many satellite-based data sets and all ozonesonde data. All records in the database are in exactly the same format which greatly facilitates processing.
- We have developed techniques for homogenizing the different sources of data but have not yet implemented them.
- We have developed methods to create a global, filled database from monthly means calculated from the BDBP
- One issue at this stage is that Greg Bodeker has no funding to continue this work and progress is likely to be slow but he is committed to supporting this activity as best as he can.