

# DATEM

## Data Archive of Tracer Experiments and Meteorology

Roland R. Draxler  
Jerome L. Heffter  
Glenn D. Rolph

NOAA Air Resources Laboratory  
1315 East West Highway  
Silver Spring, Maryland 20910

14 August 2001  
(Last Revision: 23 July 2002)

### Abstract

The availability of meteorological re-analysis archives at several international centers provides an opportunity to link high quality modern meteorological data with the data from many older long-range tracer experiments. A set of CDs and web site has been created that contains the experimental data, relevant reports, meteorological data, statistical analysis, and display software, all in a common non-proprietary format for PC or UNIX applications. This data base permits the atmospheric transport modeling community to conduct various verification and sensitivity studies and compare model results with each other on a common basis. Currently, only longer range (100's to 1000's of km downwind) experimental data are considered, consistent with the spatial and temporal resolution of the meteorological re-analysis data. The DATEM archive (software and data files) can be accessed through the ARL web page <http://www.arl.noaa.gov/datem>. User's of the data base are encouraged to share any additional analysis software that they might develop and prepare and submit additional experimental data to the archive.

### 1. Introduction

The long-range transport of pollutants in the atmosphere has received considerable attention in recent years, corresponding with a comparable number of models to address these issues. In particular there has been consistent emphasis on nuclear reactor accidents since the Chernobyl accident by the International Atomic Energy Agency (IAEA) and the World Meteorological Organization (WMO) through the organization of WMO's Regional Specialized Meteorological Centers (RSMCs). Model

development was further enhanced by the European Tracer Experiment (ETEX), which resulted in the evaluation of 47 different atmospheric transport models (ATMs). The eventual ratification of the Comprehensive Test Ban Treaty (CTBT) will require the use of ATMs to attribute measured air concentrations to a particular source location which may result in the development of a whole new class of models to address these requirements. The introduction of more fuel efficient jet aircraft engines (running at higher temperatures) has made modern commercial aircraft very sensitive to volcanic ash from eruptions. Volcanic Ash Advisory Centres (VAAC), designated by the International Civil Aviation Organization (ICAO), use ATMs to forecast the location of volcanic ash, so that aircraft may be warned and rerouted around the affected areas.

The twelfth meeting of the WMO's Commission for Atmospheric Sciences (CAS) emphasized the importance of the coordinating role of WMO in emergency response activities and the increasing number of models to address the long-range transport of pollutants. The Commission suggested that there was a need to create a data base of all known field experiments and the corresponding meteorological data in a common format that could then be used for model verification and development. The USA and Australia agreed to assemble a small expert group to consider the tasks required to address this issue and report their results to the Working Group (CAS, 1999) on Environmental Pollution and Atmospheric Chemistry. Based upon the expert group report, the US funded a small pilot study to organize some of the US sponsored experimental data into a common format. The results of this pilot study are described in this report.

Most, if not all, ATMs are linked to one or more operational meteorological forecast models or meteorological archives, which is then used by the ATM. The forecasts by the meteorological models are routinely evaluated and compared against common performance standards. Although similar verification standards have been developed for ATMs, there is a sense in the dispersion modeling community that there is very little data available to perform these evaluations. Perhaps because these data are not as easily obtained and are usually limited to single events or controlled (and expensive) field experiments, not available on a routine basis. However, there are many isolated experiments, some of which are controlled tracer releases and others from the sampling of tracers of opportunity. The problem with many of these data sets is that some are now decades old, reports are difficult to locate, and if available in electronic form, the data are in various formats, and not all experiments have archived the corresponding meteorological data.

The recent completion of meteorological re-analysis projects at several international meteorological centers, provides an opportunity to link high quality modern meteorological data with each of these dispersion experiments. This pilot study incorporated data from five long-range dispersion experiments, to be described in more detail below, which represent a collection of more than 19,000 air concentration samples, re-analysis fields from the National Center for Atmospheric Research (NCAR) / National Centers for Environmental Prediction (NCEP) re-analysis project, and statistical analysis programs based upon the ATMES-II evaluation of ETEX. All the emissions and sampling data are in space delimited text files, easily used by computer programs or imported into any

spreadsheet. Meteorological data fields are provided in a unpacked binary format, one record per variable per time period. The statistical programs are all written in Fortran and include the PC executables (using Compaq Visual Fortran 6.1) and with the source code so that the programs can be compiled on other platforms. All the data fit on 3 CDs, with the first CD containing all the experimental data, software, and meteorological data from 3 experiments. The other 2 CDs contain meteorological data from the two other experiments. The CDs can be ordered on-line, or the individual data files can be downloaded directly from the DATEM web site ( <http://www.arl.noaa.gov/datem> ). Users of the data base are encouraged to use the web site to share any additional analysis software they may develop or their own experimental data.

Although a substantial amount of data and software are provided through the pilot project, each user will need to convert their own dispersion model output to the DATEM format. The measurement data file format and the model calculation file formats are identical. The statistical analysis program matches each calculation record to a measurement record (one record per sample). Therefore each model participant will have to write a program that computes the calculated concentration corresponding with each measurement record. As an example, a Fortran program is provided that converts the binary output file from the HYSPLIT model to the DATEM format. Measurement data, meteorological data, model output files, and statistical results, all reside in their own directories with subdirectories that correspond to each experiment. In this structure, more experiments can easily be added in the future.

The meteorological data are provided in their most basic form, however, they may still need to be reformatted or interpolated prior to use in an ATM. A simple “binary unformatted” read may be sufficient to input the data in most models. Some interpolation may be required as the data fields are provided on sigma surfaces in the vertical and a Gaussian grid in the horizontal. As a convenience and to permit participation by a wider range of researchers, it is not necessary to have an ATM to evaluate these experimental data. Configuration files for HYSPLIT, an ATM for long-range applications available for download over the Internet ( <http://www.arl.noaa.gov/hysplit> ), are provided for each experiment. The meteorological re-analysis data reformatted for input to HYSPLIT and model calculations already in the DATEM format are also available.

In the following sections more detail is provided about each tracer experiment included in the archive, the meteorological data format, dispersion model calculation procedures, statistical analysis, graphical display of results, and the data distribution channels.

## **2. Experimental Data Description**

The archive consists of data from five experiments: 1) ACURATE - the Atlantic Coast Unique Regional Atmospheric Tracer Experiment from 1982 and 1983, 2) ANATEX - the Across North America Tracer Experiment from 1987, 3) CAPTEX - the Cross Appalachian Tracer Experiment

from 1983, 4) INEL74 - Idaho National Engineering Laboratory releases in 1974, and 5) OKC80 - a single tracer release from Oklahoma City in 1980. The above acronym abbreviations are used for all examples and data directory names. Detailed information on each experiment is provided in the original reports in “pdf” which can be found in the *document* directory. The descriptions below provide a brief summary of each experiment and any conversions that were performed on the data. In general, all experimental results were converted to mass units in terms of emission and air concentrations. All air concentrations have been multiplied by  $10^{+12}$  (for pico units) and include the subtraction of ambient background. Therefore, air concentrations are directly proportional to the emission rate times the atmospheric dispersion factor (supplied by the ATM). Emission and measurement data are found in the *exp\_data* directory in the appropriated experiment subdirectory with file names of *emis-t?.txt* and *meas-t?.txt*, respectively. The “?” represents a single numeric digit identifying the tracer number. Note that the experimental data placed in the DATEM archive have undergone a certain amount of reprocessing to produce more consistent data. Some of the data in the original reported data format are available at <http://www.arl.noaa.gov/ss/transport/tracer.html>

## 2.1 ACURATE Tracer #1 - Krypton<sup>85</sup>

The ACURATE (Heffter et al., 1984) experiment consisted of measuring the Kr-85 air concentrations from emissions of the Savannah River Plant, SC. Twelve and 24 hour average air concentrations were collected for 19 months (March 1982 - September 1983) at 5 locations along the United States east coast from 300 to 1000 km from the plant (Fayetteville, NC to Murray Hill, NJ). Hourly Kr-85 emissions (5 March 1982 - 30 September 1983) are given in Curies and air concentration measurements (9 March 1982 - 30 September 1983) are in pCi / m<sup>3</sup>. Ambient background concentration was subtracted from measured concentrations according to the station: FAY - 19.0, TAR - 19.0, NOR - 19.1, SAL - 19.2, MUR - 19.3. The ambient background concentration was determined to be at the point at which there was a large change in slope of the cumulative concentration distribution at each station. Background varies by latitude, increasing to the north due to the prevalence of nuclear fuel reprocessing in the northern latitudes.

## 2.2 ANATEX Tracer #1 - PTCH (GGW); #2 - PDCH (STC); #3 - PMCH (STC)

ANATEX (Draxler and Heffter, Eds, 1989) consisted of 66 PerFluorocarbon Tracer (PFT) releases (33 each from two different locations - January 5<sup>th</sup> to March 26<sup>th</sup>, 1987) every two and one half days. Air samples were collected for 3 months (January 5<sup>th</sup> to March 29<sup>th</sup>, 1987) over 24 h periods at 75 sites covering most of the eastern US and southeastern Canada. PTCH (perfluoro-trimethylcyclohexane) was released from Glasgow, Montana (GGW), and PDCH (perfluoro-dimethylcyclohexane) and PMCH (perfluoro-monomethylcyclohexane) from St. Cloud, Minnesota (STC). Release units are given in grams and air concentrations in pg / m<sup>3</sup>. Note that the PMCH releases from STC are coincident with the PDCH releases and hence do not provide any “meteorologically independent” data. The measurement data as reported already had ambient background removed. However, concentrations were reported in volume units (df/l) and were

converted to mass units according to the molecular weight of the tracer by the following conversion factors: PMCH  $1.56 \times 10^{-12}$ , PDCH  $1.79 \times 10^{-12}$ , PTCH  $2.00 \times 10^{-12}$ . All measured data have been included in the archive, regardless of any quality flags or uncertainty reported with the raw data.

### **2.3 CAPTEX**                    Tracer #1 - PMCH

The CAPTEX, (Ferber et al., 1986) during September and October of 1983, consisted of six 3 h PFT releases (September 18<sup>th</sup> to October 29<sup>th</sup>, 1983), four from Dayton, Ohio and two from Sudbury, Ontario, Canada with samples collected at 84 sites, 300 to 800 km from the source, at 3 h and 6 h averages for about a 48 hour duration after each release (September 18<sup>th</sup> to October 30<sup>th</sup>, 1983). One additional short (30 min) tracer release from Dayton was not evident in the sampling data. Concentrations were already reported as above background but were converted from volume (fl/l) to mass units ( $\text{pg} / \text{m}^3$ ) by the relation  $15.6 \times 10^{-12}$ .

### **2.4 INEL74**                    Tracer #1 - Krypton<sup>85</sup>

The INEL74 experiment (Ferber et al., 1977; Draxler, 1982) consisted of a little over two months of Kr-85 releases (February 27<sup>th</sup> to May 5<sup>th</sup>, 1974) from Idaho and continuous 12 h sampling (February 27<sup>th</sup> to May 4<sup>th</sup>, 1974) at 11 locations in a line about 1500 km downwind (Oklahoma City, OK to Minneapolis, MN). Emissions are given in Curies and air concentrations in  $\text{pCi} / \text{m}^3$ . The same ambient background concentration ( $13.7 \text{ pCi} / \text{m}^3$ ) was subtracted from all stations

### **2.5 OKC80**                    Tracer #1 - PMCH      Tracer #2 - PDCH

The OKC80 experiment (Ferber et al., 1981) consisted of single release of two different PFT tracers (July 8<sup>th</sup>, 1980) over a 3 hour duration with samples of 3 hour duration collected at 10 sites 100 km and 35 sites 600 km downwind from the Oklahoma City release point from July 8<sup>th</sup> to July 11<sup>th</sup>, 1980. Emissions are given in grams and air concentrations in pico-grams /  $\text{m}^3$ . Concentrations were reported in volume units (df/l) and were converted to mass units according to the molecular weight of the tracer by the following conversion factors: PMCH  $1.56 \times 10^{-12}$  and PDCH  $1.79 \times 10^{-12}$ . Background concentrations of 2.4 fl/l and 30.0 fl/l were subtracted for PMCH and PDCH, respectively.

### **2.6 Emission Data File Format**

Emission information is found in the `\exp_data` directory under the specific `\experiment` subdirectory. Emission files are in an ASCII text format, the first record identifying the directory and file name, the second line identifying all the variables that follow on all subsequent records, with one record for each emission. All times are given in UTC (GMT). The file names all start with `emit-t?.txt`. The “?” is the tracer number as defined in the above sections for each experiment. The emission data records have the following format:

Year	I4	Four digit year for emission start
Month	I3	Two digit month for emission start
Day	I3	Two digit day for emission start
Hour-Minutes	I5	Four digit hour-minutes for emission start
Duration	I5	Four digit hour-minutes for duration of the emission
Latitude	F6.2	Latitude of the emission point (Positive is north)
Longitude	F8.2	Longitude of the emission point (Positive is east)
Emission	F8.0	Total emission in mass units over the duration

## 2.7 Sampling Data File Format

Measurement data are found in the `\exp_data` directory under the specific `\experiment` subdirectory. Sampling data files are in an ASCII text format, the first record identifying the directory and file name, the second line identifying all the variables that follow on all subsequent records, with one record for each sample collected. Samples, although generally in chronological order, may be organized differently with each experiment. All times are given in UTC. The file names all start with `meas-t?.txt`. The “?” is the tracer number as defined in the above sections for each experiment. The sampling data records have the following format:

Year	I4	Four digit year for sample start
Month	I3	Two digit month for sample start
Day	I3	Two digit day for sample start
Hour-Minutes	I5	Four digit hour-minutes for sample start
Duration	I5	Four digit hour-minutes for duration of the sample
Latitude	F6.2	Latitude of the sampling point (Positive is north)
Longitude	F8.2	Longitude of the sampling point (Positive is east)
Concentration	F8.1	Average air concentration in (pico) mass units per cubic meter
Site Name	A8	Character string for site identification

## 3. Meteorological Re-analysis Data

The NCEP/NCAR Re-analysis Project (Kalnay et al., 1996) is a joint project between NCEP and NCAR. The goal of this joint effort was to produce new atmospheric analyses using historical data (1958 - 1997) and analyses of the current atmospheric state (Climate Data Assimilation System - CDAS). The data assimilation and the model used are identical to the global system implemented operationally at NCEP on 11 January 1995, except that the horizontal resolution is T62 (about 210 km). The data base has been enhanced with many sources of observations not available in real time for operations, provided by different countries and organizations. For the DATEM archive, the meteorological data fields (surface pressure, temperature, humidity, horizontal wind components, and vertical velocity) were processed on sigma levels for use by the ATMs, rather than the more common mandatory pressure surfaces. The sigma level data provide more detail within the boundary layer.

Meteorological data files can be found in the `\met_data` directory, by experiment name subdirectory. Files contain data for an entire month, regardless of the length of the experiment. File names follow the convention: `RA{year}{month}.bin`, where “year” is the four digit year and “month” is the two digit month. RA is the abbreviation for Re-analysis Archive. File sizes are about 60 Mb per month.

### 3.1 Archive Processing

The spectral coefficients on 28 model sigma surfaces were processed to obtain the fields 4 times per day on a global Gaussian grid of 1.875 deg resolution. A 41 x 21 regional sub-grid covering the continental US and Canada was extracted from the global files with the southwest corner at 21.904N 127.500W and the northeast corner at 59.999N 52.500W. The data were written as unformatted binary (Big-Endian) records, one month per file. Each file is composed of one header record of 80 characters, followed by 6 data records per time period, one record per variable. Each variable contains all the grid elements of rank 2 for surface pressure and rank 3 for all the other variables. The left most index varies first. The data elements in each array vary from west to east, south to north, and bottom to top, with increasing index values. All meteorological variables are four byte real. The variable on a record is preceded by two four-byte integers, indicating the valid day and hour of the data record. If difficulties are experienced reading these data records, note that each unformatted record's length is the length of the data in bytes plus 8. An unformatted data record is preceded and trailed by a 4 byte integer that gives the length of the data record in bytes. Six data variables are given in each file: surface pressure (KPa), air temperature (K), west to east component of the wind (m/s), south to north component of the wind (m/s), pressure vertical velocity (Pa/s). Each record contains one variable for a specific time: day hour variable. Data are written in the following order:

	Record 1 - 80 characters: the data file name (year & month) and variables
Time 1	Record 2 - Day, Hour, Surface Pressure
	Record 3 - Day, Hour, Air Temperature
	Record 4 - Day, Hour, Specific Humidity
	Record 5 - Day, Hour, West to east wind component
	Record 6 - Day, Hour, South to north wind component
	Record 7 - Day, Hour, Pressure vertical velocity
Time 2	Record 8 - Day, Hour, Surface Pressure
	...
End-of-Month	

Details about the vertical sigma surfaces, latitude, and longitude points that correspond with each grid location can be found in Appendix A. Vertical sigma levels are defined as  $(P - P_{top}) / (P_{sfc} - P_{top})$ , where P is pressure at the “top” of the model and the “sfc” surface. An example Fortran program “`readbin`” is provided in the `\met_prm` directory which reads each data record and writes the value of the variable at the grid center to standard output. This program can easily be incorporated into other programs (or ATMs) to read the binary meteorological data files. The PC executable can be found in

the *\Release* subdirectory.

### 3.2 HYSPLIT Compatible Meteorological Data

As a convenience for those without their own ATM and to provide researchers with results from another ATM to compare with their own calculations, meteorological data are also provided in a HYSPLIT compatible format. The DATEM CDs can be ordered with meteorological data in standard binary format (*.bin* files in the *\met\_data* directory) as noted in the above paragraph or as HYSPLIT compatible packed data (*.arl* files in the *\met\_data* directory). Otherwise the file and directory naming convention are identical. Detailed information on the HYSPLIT compatible format can be found in Chapter 5 of the HYSPLIT User's Guide (Draxler, 1999). The ARL packed meteorological data files are directly usable by either the PC or UNIX versions of HYSPLIT.

## 4. Dispersion Model Simulations

Each ATM will have its own specific configuration requirements. A complete simulation for all experiments requires almost 19,000 hours of model time, and is clearly dominated by the 19 month duration ACURATE. However, it may be of interest to understand how one particular ATM (HYSPLIT) was run for each experiment. The model's configuration files for each experiment can be found in the *\mdl\_cfg* directory under specific experiment subdirectories. Each experiment has one or more *CONTROL* files that are read during the model's startup procedure and a "namelist" file, *SETUP.CFG*, that sets additional simulation options. Reference to the User's Guide is essential to understand the implications of modifying these files. Some explanation can be found in each directory's *Readme.txt* file. MS-Windows shortcut files (*Run.lnk*) and Batch files are provided to illustrate how the model can be started from the command line or in batch mode rather than through the GUI. These files may have to be edited to correspond with the directory structure on specific PC or UNIX installations.

During the simulation, the model's binary concentration output files are written to the *\mdl\_data* directory under the appropriate experiment subdirectory. It is suggested that this convention be followed as well with other ATMs, because the format conversion programs, statistical analysis, and display programs look for the data files under this particular directory structure and file naming convention.

### 4.1 ACURATE

Due to the long duration of the experiment, the simulation was divided into seven individual runs. Each run (except the first) was started 5 days prior to the start of the simulation period to permit the tracer to build up over the computational grid. The final text output file was the result of the merging the individual seasonal simulation files. Although the sampling duration was nominally 12 hours, it was



necessary to produce model outputs every hour because the sampling data would shift by one hour when local time changed between daylight savings and standard time. Sampling station operators worked on local time. As noted earlier, all times in the archive have been converted to UTC.

Concentration output grid resolution:	50 km / 100 m / 1 h
Grid domain to covers samplers:	30N-45N and 70W-85W
Simulation periods	P1: March, April, May 1982 = 2112 h
	P2: June, July, August 1982 = 2328 h
	P3: September, October, November, 1982 = 2304 h
	P4: December, January, February, 1983 = 2280 h
	P5: March, April, May, 1983 = 2328 h
	P6: June, July, August, 1983 = 2328 h
	P7: September, 1983 = 864 h

#### 4.2 ANATEX

Concentration output grid resolution:	50 km / 100 m / 24 h
Grid domain to cover samplers:	Centered 40.0N 85.0W spanning 35 Lat 55 Long
Simulation periods:	2040 hours starting 1700 UTC January 5 <sup>th</sup> , 1987

#### 4.3 CAPTEX

Concentration output grid resolution:	50 km / 100 m / 3 h
Grid domain to cover samplers:	Centered 42.0N 78.0W spanning 10 Lat 15 Long
Simulation periods:	1032 hours starting 1700 UTC September 18 <sup>th</sup> , 1983

#### 4.4 INEL74

Concentration output grid resolution:	50 km / 100 m / 2 h
Grid domain to cover samplers:	Centered 40.0N 95.0W spanning 15 Lat 10 Long
Simulation periods:	1608 hours starting 0000 UTC February 27 <sup>th</sup> , 1974

#### 4.5 OKC80

Concentration output grid resolution:	20 km / 100 m / 15 min
Grid domain to cover samplers:	Centered 37.0N 95.0W spanning 10 Lat 10 Long
Simulation periods:	73 hours starting 1900 UTC July 8 <sup>th</sup> , 1980

#### 4.6 Model Simulation Output Format Specifications

Model simulation results should be written to the *\mdl\_data* directory under the specific *\experiment* subdirectory. Data files should be in an ASCII text format, identical to the measurement data, the first record identifying the directory and file name, the second line identifies all the variables

that follow on all subsequent records, with one record for each sample collected. All times are given in UTC. The file names follow the convention *ct?\_001.txt*. The “?” is the tracer number as defined for that experiment and 001 (002, 003, etc) would represent different model results, simulation variations, or any user defined character string. The model output data records have the following format:

Year	I4	Four digit year for sample start
Month	I3	Two digit month for sample start
Day	I3	Two digit day for sample start
Hour-Minutes	I5	Four digit hour-minutes for sample start
Duration	I5	Four digit hour-minutes for duration of the sample
Latitude	F6.2	Latitude of the sampling point (Positive is north)
Longitude	F8.2	Longitude of the sampling point (Positive is east)
Concentration	F8.1	Average air concentration in (pico) mass units per cubic meter
Site Name	A8	Character string for site identification

Most ATMs are not designed to easily output air concentration data at various locations and sampling intervals. The statistical programs require the model predictions to correspond with each measurement. One suggestion is to create a program to convert the normal ATM output format to the DATEM compatible format. An example program is available for HYSPLIT simulations as discussed in the next section. This program could be modified to handle output from other ATMs.

#### 4.7 HYSPLIT Output Conversion Program Example

The program “*c2datem*”, which converts HYSPLIT binary air concentration output files to DATEM format, can be found in the `\stat_pgm\c2datem` directory (executable in *Release*). The program can be used as a basis to develop conversion programs for other ATMs. The program reads all of the experimental data into memory and then sequentially reads the model’s concentration output file. Each model output time period, all sampling data are scanned to determine if the time duration of the calculation falls within the duration of one of the samples. In that situation, the calculated air concentration is accumulated for that station. If the calculation time falls at the end of a sampling period, the average concentration is computed for that station and the output is written to a text file in identical format to the measured data. The program logic permits the ATM calculations to be matched to the measurement data, regardless of the organization of the data in the file. The measurement data do not all need to have the same sampling intervals. The only requirement is that the model’s output interval is less than or equal to the measurement sampling interval. For instance the model’s output interval for INEL74 was set to 2 hours to accommodate that 10 h duration samples were collected starting at 1000 and 0200. The command line options for the program are as follows:

```
c2datem -i -o -m -c -x -z -p -d
-i[input concentration file name]
-o[output concentration file name]
-m[measurement data file name]
-c[input to output concentration multiplier]
```

-x[n(neighbor) or i(interpolation)]  
-z[level index]  
-p[pollutant index]  
-d[write to diagnostic file: c2datem.txt]

As an example, to convert HYSPLIT output from the OKC80 experiment:

```
c2datem -ict1.bin -oct1.txt -m..\exp_data\okc80\meas-t1.txt -xi -c1.0E+12
```

where *ct1.bin* defines the HYSPLIT output file, *ct1.txt* is the model calculation output interpolated to the sampling stations in a text format, with model concentrations multiplied by  $10^{12}$  before output. Linear interpolation (-xi) is used compute station values from the model's gridded concentration field. The pollutant number and output level default to 1 (See the HYSPLIT User's Guide for more information on these parameters).

## 5. Statistical Analysis

The model evaluation protocol was designed following the procedures used by Mosca et al. (1997) for the ATMES-II study and Stohl et al. (1998) in evaluating several of the same tracer experiment data presented here. Reviewing these previous studies, four broad categories were defined for a model's evaluation: 1) the scatter among paired measured and calculated values, 2) the bias of the calculation in terms of over- or under-calculation, 3) the spatial distribution of the calculation relative to the measurements, and 4) differences between the distribution of unpaired measured and calculated values. In each category, one or more evaluation parameters are defined, such as the percent over-prediction, number of calculations within a factor of 2 and 5 of the measurements, normalized mean square error, correlation coefficient, bias, fractional bias, figure of merit in space, and the Kolomogorov-Smirnov parameter representing the differences in cumulative distributions. All of these are defined in more detail in Appendix B.

Both Mosca and Stohl recognized the problem in dealing with the uncertainties of "near background" measurement data and avoiding statistical parameters that may be sensitive to small variations in the measurement values such as ratios between measured and calculated concentration. Each had their own approach to the meaning of a "zero" measurement. Stohl using a Monte-Carlo approach to add and subtract small values from the measurement to test the robustness of the statistical parameters, while Mosca used a 6 hour filter before and after the arrival of the plume at each station to select samples for analysis. Both defined threshold measurement values, below which concentrations were assumed to be zero. A hybrid method, independent of a specific experiment's detection limits and to include all the data, was designed for this study. The user may select a zero threshold probability, that is a cumulative probability level, below which measured concentrations are assumed to be zero. For instance, with a selection of 10%, the analysis might show that 10% of the measured concentrations fall below say "5 units" and 90% are greater. Therefore in all subsequent analyses, measured concentrations below "5 units" are treated as zero. A command line switch in the analysis program permits the inclusion or exclusion of "0-0" measured-calculated pairs. Because the treatment

of zero's can influence which data pairs are analyzed, the suggested default approach for model comparison purposes is to use all the data and rely upon more robust statistical parameters and avoid any ratio dependent statistics. For instance, the ratio of the mean predicted over mean measured concentration is devoid of problems, while the mean of the ratio of predicted over measured concentration is fraught with difficulty. If problems are suspected due to the treatment of zeros, either temporal or spatial averaging of the data is recommended, in this way all the data are used in the analysis.

For evaluation comparison purposes it would be desirable to have a single robust parameter which could be used to determine an overall degree of model performance. Stohl evaluated many of the above parameters and found that the ratio based ones (such as FA2 and FA5) were the most sensitive to measurement errors while the correlation coefficient was one of the most robust. For the DATEM metric, one statistical measure that could be used to represent the overall model performance, it was decided to take the sum of one parameter from each of the four model evaluation categories (correlation, fractional bias, figure of merit, and Kolomogorov-Smirnov) which could be normalized to the range of 0-1 (worst to best).

Model analyses can be applied in a *global analysis*, where all pairs are treated individually, a *space analysis* which is the result of temporally averaging all the data so that only spatial variability remains, and a *time analysis* which is the result of averaging all the observations at each time so that only temporal variability remains.

## 5.1 Statistical Program Command Line Options

The statistical analysis program source code can be found in `\stat_pgm\statmain` and the executable in `\stat_pgm\statmain\Release`. However the program should be run from the `\mdl_stat` experiment subdirectory where a shortcut link already exists with the appropriate command line options. There is a shortcut for the global calculation using all data pairs (`stat_gbl.lnk`) and one for the spatially averaged calculation (`stat_avg.lnk`). In addition, each subdirectory has a batch file replica for the global execution. The program needs to reside in the `\mdl_stat` directory because on the command line only the experiment name, model variation, and tracer number are required. Measurement data and model prediction files are then opened in the appropriate `\exp_data` and `\mdl_data` directories following the tracer number and model variation naming convention. The command line options are summarized below:

```
statmain -d -t -m -z -a -o -x  
-d[directory name of experiment]  
-t[tracer number: 1,2,3, etc]  
-m[model variation string]  
-z[percentile level for zero measured]  
-a[space (s) or time (t) averaging]  
-o[creates merged output file]  
-x[exclude 0,0 pairs from statistics]
```

and where the statistical results will be written to a file named: *{gbl/avg}{tracer#}\_{variation}.txt*. Normally the results file starts with “*gbl*” and represents a global statistic that uses all the data. However setting the averaging flag to either *-as* or *-at* will result in spatial or temporal averaging and an output file name starting with “*avg*”. Setting the *-o* flag causes a text output file (starting with “*dat*”) to be written that contains both the measurement and prediction data on the same line. This file can then be imported into other applications (spreadsheets or GIS). A command line example for ACURATE with the minimum number of options would be: `statmain -dacurate -m001 -t1`

### 5.3 Explanation of Sample Output

The sample statistical output from ETEX is shown in Appendix C. Although the data from that experiment are currently not included in the DATEM archive, it is of interest to compare the DATEM statistical results to those published by Mosca in the ATMES-II evaluation (model 120 page 203) for the same model (HYSPLIT) using the same meteorological data from ECMWF. An exact comparison is not expected as the ATMES-II evaluation applied several filters to the data that are not used in DATEM. In addition, the model has undergone several changes since the ETEX evaluation. Results for several of the primary statistical parameters that can be compared are as follows:

<u>Statistic</u>	<u>DATEM</u>	<u>ATMES-II</u>
Correlation	0.48	0.44
NMSE	23.25	23.6
Bias	37.98	180
FMS	57.27	61
FA2	15.95	35
FA5	34.64	43

Some of the differences are because the DATEM statistics include all measured predicted pairs and apply no special rules for zeros. For reference, HYSPLIT statistical results are available in each *\mdl\_stat* experiment directory (*gbl1\_001.txt* and *avg1\_001.txt*). The global and spatially averaged statistical results are also summarized in Appendix D for each of the experiments.

### 5.4 Graphical Output Display

The model predictions and measurements can be displayed for any sampling period by using the merged output file created by the *statmain* program with the *-o* flag set as noted in Section 5.2. Software is provided in the *\plot\_pgm\c2grads* directory to convert the *dat{tracer#}\_{variation}.txt* file in the *\mdl\_stat{experiment}* directory to a GrADS (Grid Analysis and Display System) formatted input file. GrADS is distributed free of charge (<http://grads.iges.org/grads>). An “X” server is also required. Versions of GrADS are available for several flavors of UNIX, Linux, PCs running Windows 95/NT or DOS, and MacIntosh computers. The graphics conversion should be run in the *\plot\_pgm* directory using the following command line options:

**c2grads** -d[directory] -m[variation] -t[tracer]

and where the command line options have the same meaning as in the *statmain* program. The GrADS conversion program reads the *dat{}* file in the experimental directory and will create four output files in the *\plot\_pgm* directory:

*merged.dat* - the input data file converted to GrADS format  
*merged.ctl* - the control file for *merged.dat*  
*merged.gs* - a script file to make a sample plot  
*dummy.ctl* - a dummy control file to create the grid for interpolation

Because the model predictions and measurement data are treated as “station data”, before proceeding it is necessary to run the GrADS program *stnmap* from the command line:

```
stnmap -i merged.ctl
```

At this point GrADS can be started and then from the GrADS command line:

```
ga-> exec makeplot.txt
```

runs the example program to display the ANATEX measured and calculated concentrations for the 24 hour sample starting January 9, 1987 (the 5<sup>th</sup> day after the experiment start). The measured concentrations are shown in Appendix E as the shaded area with an overlay of the model calculation contours. Other days can be plotted by simply changing the time, for instance for the next day’s sample: “set t 6”. Additional detail on how to use GrADS is beyond the scope of this document. However, the scripts and control files are provided as a starting point to simplify the development of other applications. More information can be found in the *\plot\_pgm\Readme.txt*, including all the files required to run this example. Data from other experiments, or with varying averaging options, can be displayed by using the *statmain* program is to create the appropriate merged output file.

## 6. Data and Software Distribution Procedures

The DATEM archive (software and data files) can be accessed through the ARL web page <http://www.arl.noaa.gov/datem>. The data files can be downloaded individually through the web page or it can be ordered as a set of 3 CD ROMS. The CD directory and file structure is outlined in Appendix F. The CDs are available with the meteorological data in either “*arl*” or “*bin*” format. Users of the DATEM archive are encouraged to add additional experimental data to the archive, preferably by hosting the data on their own web sites. The appropriate links will be added to the DATEM web site. User’s are also encouraged to add their analysis programs or suggest modifications to the existing programs. An area of the DATEM site will be reserved for software contributions.

## 7. Acknowledgements

This project was funded by the U.S. Air Force Technical Applications Center. Our thanks to Julian Wang, who provided conversion software and consulting advice in using the NCEP reanalysis data archives.

## 8. References

Draxler, R.R., 1999, HYSPLIT\_4 User's Guide, NOAA Tech. Memo., ERL ARL-230, 35 p.

Draxler, R.R., 1982, Measuring and modeling the transport and dispersion of Kr-85 1500 km from a point source, Atm. Environ., *16*: 2763-2776.

Draxler, R.R. and J.L. Heffter (Eds), 1989, Across North America Tracer Experiment (ANATEX) Volume I: Description, Ground-Level Sampling at Primary Sites, and Meteorology, January, NOAA Tech Memo ERL ARL-167.

Ferber, G.J., J.L. Heffter, R.R. Draxler, R.J. Lagomarsino, F.L. Thomas, R.N. Dietz, 1986, Cross-Appalachian Tracer Experiment (CAPTEX '83) Final Report, January, NOAA Tech Memo ERL ARL-142.

Ferber, G.J., K. Telegadas, J.L. Heffter, C.R. Dickson, R.N. Dietz, P.W. Krey, 1981, Demonstration of a Long-Range Tracer System using Perfluorocarbons, Final Report, January, Tech. Report EPA-600.

Ferber, G.J., K. Telegadas, J.L. Heffter, M.E. Smith, 1977, Air Concentrations of Krypton-85 in the Midwest United States during January-May 1974, Atm. Environ. *11*, 379-385.

Heffter, J.L., J.F. Schubert, and G.A. Meade, 1984, Atlantic Coast Unique Regional Atmospheric Tracer Experiment, October, NOAA Tech Memo. ERL ARL-130.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996, the NCEP / NCAR 40-Year Reanalysis Project, Bull Am. Meteorol. Soc. *77*, 437-471.

Mosca, S., R. Bianconi, R. Bellasio, G. Graziani, and W. Klug, 1998, ATMES II - Evaluation of long-range dispersion models using data of the 1<sup>st</sup> ETEX release, European Commission Joint Research Centre, EUR 17756 EN, 459 p + Appendicies, Office for Official Publications of the European Communities, L-2985, Luxembourg

Stohl, A., M. Hittenberger, and G. Wotawa, 1998, Validation of the Lagrangian particle dispersion model Flexpart against large-scale tracer experiment data, Atm. Environ., *32*: 4245-4264.

CAS, 1999, Report from the Ad Hoc Group of Experts (R.R. Draxler, G.D. Hess, R. Maryon, G. Graziani, W. Klug, F. Gifford) on the Long-range Transport and Dispersion Model Verification Data Base, submitted to the CAS Expert Panel, March 1999, Zurich, CAS Working Group on Environmental Pollution and Atmospheric Chemistry.



## APPENDIX A - Grid Point Positions for the Unpacked Binary Data

	Latitude South to North	Longitude West to East	Sigma Interface Bot to Top	Sigma Center Bot to Top
	-90 -> +90	0 -> 360	1 -> 0	1 -> 0
1	21.904	232.500	1.000	.995
2	23.809	234.375	.990	.982
3	25.714	236.250	.974	.964
4	27.619	238.125	.955	.943
5	29.523	240.000	.931	.916
6	31.428	241.875	.901	.884
7	33.333	243.750	.866	.846
8	35.238	245.625	.825	.801
9	37.142	247.500	.778	.751
10	39.047	249.375	.724	.694
11	40.952	251.250	.665	.633
12	42.856	253.125	.601	.568
13	44.761	255.000	.535	.502
14	46.666	256.875	.469	.436
15	48.571	258.750	.403	.372
16	50.475	260.625	.341	.312
17	52.380	262.500	.284	.258
18	54.285	264.375	.233	.210
19	56.189	266.250	.188	.168
20	58.094	268.125	.149	.133
21	59.999	270.000	.117	.103
22		271.875	.089	.078
23		273.750	.067	.058
24		275.625	.049	.042
25		277.500	.035	.029
26		279.375	.023	.018
27		281.250	.014	.010
28		283.125	.007	.003
29		285.000	.000	
30		286.875		
31		288.750		
32		290.625		
33		292.500		
34		294.375		
35		296.250		
36		298.125		
37		300.000		
38		301.875		
39		303.750		
40		305.625		
41		307.500		

## APPENDIX B - Statistical Analysis Equations

### B.1 Scatter Diagram

The scatter between measured and calculated values can be defined in several different ways. The simplest approach is to calculate the percent of over-predictions:

$$FOEX = \left[ \frac{N_{(Pi > Mi)}}{N} - 0.5 \right] \times 100$$

where N in the numerator is the number of pairs where the prediction exceeds the measurement and the N in the denominator is the total number of pairs in the evaluation. FOEX can range from -50% to +50%. A more restrictive test would be the percent of predictions that are within a factor of two or five of the measurement:

$$FAa = \left[ \frac{N_{(y-y0=[x-x0]a)}}{N} \right] \times 100$$

where a = 2 or 5. In the calculation of FOEX and FAa, all 0-0 pairs are always excluded. Another common statistic is the normalized mean square error in which the average of the square of the differences is divided by the product of the means:

$$NMSE = \frac{1}{NPM} \sum_i (Pi - Mi)^2$$

which can be very sensitive to differences in model predictions and measured concentration distributions. The linear correlation coefficient is perhaps the most common statistic:

$$R = \frac{\sum_i (Mi - \bar{M}) \cdot (Pi - \bar{P})}{\sqrt{\sum (Mi - \bar{M})^2} \sqrt{\sum (Pi - \bar{P})^2}}$$

with values that range between -1 and +1.

### B.2 Bias

The mean difference between paired predictions and measurements yields a measure of the prediction bias in concentration units:

$$\bar{B} = \frac{1}{N} \sum_i (P_i - M_i)$$

However it is necessary to determine if the bias is significantly different from zero. This is evaluated from the Confidence Interval which is defined from the bias:

$$CI = \pm Ta \sqrt{\frac{1}{N(N-1)} \sum_i (B_i - \bar{B})^2}$$

where Ta is the student t-distribution value at probability a. The computational default is to define the probability at the 99% level so that if the CI includes "0", then the bias is not significantly different from zero. The fractional bias, a normalized measure of bias, has also been used by many authors:

$$FB = 2\bar{B} / (\bar{P} + \bar{M})$$

and where the FB can range from -2.0 to +2.0

### B.3 Spatial Distribution

The Figure of Merit in Space is defined as the percentage of overlap between measured and predicted areas:

$$FMS = 100 \frac{N(p > 0) \cap N(m > 0)}{N(p > 0) \cup N(m > 0)}$$

the intersection over the union of predicted and measured concentrations. Mosca defined the FMS using area for a fixed sampling time and concentration level after contouring both the measured and predicted concentrations. Similar to the approach used here, Stohl simply defined the FMS to represent the sets of model results and measurement data exceeding a predefined threshold.

### B.4 Cumulative Distributions

The Kolomogorov-Smirnov parameter is defined as the maximum difference between two cumulative distributions:

$$KS = \text{Max} |C(M_k) - C(P_k)|$$

where C is the cumulative distribution of the measured and predicted concentrations over the range of k values such that C is the probability that the concentration will not exceed  $M_k$  or  $P_k$ . The KS parameter is important in that it is the only statistic that is computed from unpaired data. It is a measure of how well the model reproduces the measured concentration distribution regardless of when or where

it occurred. The maximum difference between any two distributions cannot be more than 100%.

### **B.5 Model Ranking**

The final model performance ranking is defined as the sum:

$$RANK = R^2 + (1 - |FB / 2|) + FMS / 100 + (1 - KS / 100)$$

where  $|R|$  represents the scatter;  $1 - |FB/2|$  the bias;  $FMS/100$  the spatial coverage; and  $1 - KS/100$  the unpaired distribution. The best model ranking result would be 4.0.

## Appendix C - Example global statistical output from ETEX

etex results file: gbl1\_001.txt

```

-----
3105      Unaveraged data points for processing
0.00      Percentile input for zero measured
0.00      Zero measured concentration value

0.48      Correlation coefficient (P=99%)
30.46     T-value (|Slope|/Standard Error)
95.84     Average measured concentration
133.82    Average calculated concentration
1.40      Ratio of calculated/measured
23.25     Normalized mean square error
3105      Number of pairs analyzed

37.98     Average bias [(C-M)/N]
12.75     Lo 99 % confidence interval
63.20     Hi 99 % confidence interval
0.33      Fractional bias [2B/(C+M)]

57.27     Fig of merit in space (%)

8.83      Factor exceeding [N(C>M)/N-0.5]
15.95     Percent C/M±2
34.64     Percent C/M±5
57.27     Percent M>0 and C>0
12.54     Percent M>0 and C=0
30.19     Percent M=0 and C>0

490.00    Measured 95-th percentile
200.00    Measured 90-th percentile
20.00     Measured 75-th percentile
0.00      Measured 50-th percentile
655.90    Calculated 95-th percentile
278.10    Calculated 90-th percentile
13.00     Calculated 75-th percentile
0.00      Calculated 50-th percentile

68.00     Kolmogorov-Smirnov Parameter

1.96      Final rank (C,FB,FMS,KSP)

```

## APPENDIX D - Summary of HYSPLIT results

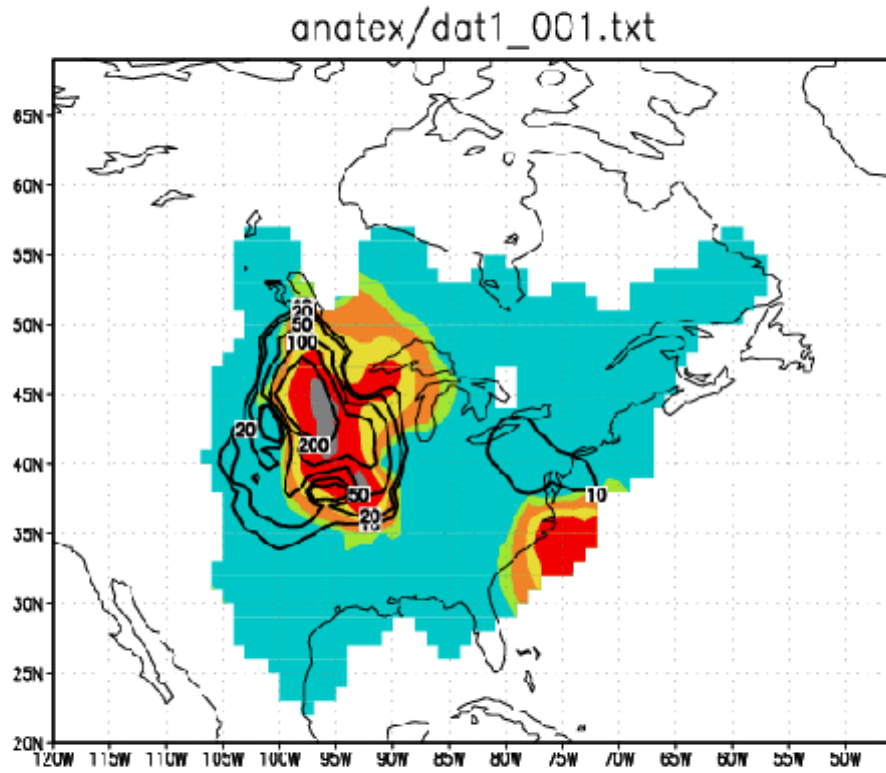
### Global

<u>Experiment</u>	<u>Data</u>	<u>R</u>	<u>FB</u>	<u>FMS</u>	<u>KS</u>	<u>RANK</u>
ACURATE	3829	0.28	-0.02	18.59	79.0	1.46
ANATEX1	5600	0.05	-0.14	34.98	59.0	1.69
ANATEX2	5482	0.10	+0.16	37.02	56.0	1.74
CAPTEX	2281	0.01	-0.17	15.34	71.0	1.36
INEL74	1269	0.00	+0.13	7.97	92.0	0.48
OKC80	697	0.16	-0.67	51.15	43.0	1.77

### Temporally Averaged

<u>Experiment</u>	<u>Data</u>	<u>R</u>	<u>FB</u>	<u>FMS</u>	<u>KS</u>	<u>RANK</u>
ACURATE	5	0.91	-0.03	100.0	39.0	3.43
ANATEX1	76	0.35	-0.29	100.0	29.0	2.69
ANATEX2	75	0.54	+0.12	100.0	31.0	2.92
CAPTEX	80	0.78	-0.18	93.59	10.0	3.35
INEL74	11	0.13	+1.39	100.0	89.0	1.43
OKC80	45	0.96	-0.65	90.70	14.0	3.36

# APPENDIX E - Example Graphical Output from ANATEX



GRADS: COLA/GES

200 -08-15-09:31

## APPENDIX F - CD Directory and File Structure

---

The archive is spread over 3 CDs. CD1 contains nine directories with all experimental data and meteorological data from three experiments: ANATEX, INEL74, and OKC80. The meteorological data from the other two experiments, ACURATE and CAPTEX, are on the remaining CDs. The directory structure is as follows:

**document** (CD1: all experiments)

Copies of PDF reports for each experiment as prepared by the experiment authors. Summarizes many details about the methodology and usually tabulates the raw data prior to any DATEM processing.

**exp\_data** (CD1: all experiments)

The emissions and sampling data processed into the standard DATEM format. Standard file name convention for each experimental subdirectory is to name the files emit-t?.txt and meas-t?.txt, respectively.

**mdl\_cfg** (CD1: all experiments)

Example simulation CONTROL, SETUP, and execution files are provided for the HYSPLIT ATM. These files can be used as guidance for each user to set up their own ATM inputs, or run their own simulations using HYSPLIT.

**mdl\_data** (CD1: all experiments)

ATM output files are written to this directory for each experiment. Execution examples are provided that convert HYSPLIT output files to the DATEM format with standard file names such as ct1\_001.txt. Users should place their ATM outputs in this directory.

**mdl\_stat** (CD1: all experiments)

Execution examples are provided that read the DATEM format experimental data (\exp\_data) and ATM results (mdl\_data) and created the statistical analysis output files (avg1\_001.txt and gbl1\_001.txt) in each experiment subdirectory.

**met\_data** (CD1: ANATEX, INEL74, OKC80) (CD2: ACURATE) (CD3: ACURATE, CAPTEX)

Meteorological data are provided at a minimum of one month increments in each experimental directory. File name ending with .bin are in standard unpacked binary format, while those ending in .arl are compatible for direct HYSPLIT input.



**met\_pgm** (CD1: all experiments)

The "readbin" directory contains a sample Fortran 90 program to read the binary meteorological data files. This program can easily be modified or incorporated into other ATMs to read the data or convert them to another format.

**plot\_pgm** (CD1)

The "c2grads" directory contains a Fortran 90 program that will convert the merged measured and predicted files output from the "statmain" program to GrADS format. In addition, the program also creates all required GrADS control files. GrADS can then be used to contour the measured and model results.

**stat\_pgm** (CD1)

The "c2datem" program is an example to convert HYSPLIT output files to the DATEM format. The "statmain" program is the primary statistical program used to analyze ATM performance.