



**A Program for the Statistical Evaluation
of Point Source Dispersion Models
Using ASTM D 6589**

**A Program for the Statistical Evaluation
of Point Source Dispersion Models
Using ASTM D 6589**

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air Quality Planning and Standards
Emissions, Monitoring, and Analysis Division
Research Triangle Park, North Carolina 27711

August 2003

DISCLAIMER

The information in this document is being reviewed in its entirety by the U.S. Environmental Protection Agency (EPA), for approval for publication as an EPA document. Mention of trade names, products, or services does not convey, and should not be interpreted as conveying official EPA approval, endorsement, or recommendation.

PREFACE

The American Society for Testing and Materials (ASTM) has published a Standard Guide D 6589 for Statistical Evaluation of Atmospheric Dispersion Model Performance. Within the Annex to this ASTM Guide, a procedure is outlined for assessing the performance of atmospheric transport and diffusion models to predict the average maximum "centerline" concentration values of a chemical species that has been released from a point source.

The EPA has developed a FORTRAN program called "Design 2" that implements the procedure documented within the Annex to ASTM Standard Guide D 6589, for the statistical evaluation of the performance of transport and diffusion models to simulate the average "centerline" concentration values.

This report describes the input files and formats of these input files for running Design 2. The report describes the information provided by the various output files created by Design 2. The report does not describe the actual statistical procedures of Standard Guide D 6589, as these are documented in detail within this Guide's Annex.

ACKNOWLEDGMENTS

This research described in this report was funded wholly or in part by the United States Environmental Protection Agency (EPA) through EPA Contract NO. 68-D-98-006, Work Assignment NO. 4-02 with EC/R Inc., who subcontracted with the Pacific Environmental Services, Inc, a MACTEC Company (PES). The EPA Work Assignment Officer was John S. Irwin. The EC/R Project Leader was Steve Fudge. The PES Project Leader was James Paumier.

TABLE OF CONTENTS

	<u>Page</u>
Preface	iii
Acknowledgments	iv
1. INTRODUCTION	1
2. OBSERVED CONCENTRATION DATA	4
3. MODELED CONCENTRATION DATA	8
4. REGIMES	10
5. CONTROL FILE	13
6. RESULTS	17
7. REFERENCES	23
APPENDIX - DIRECTORY STRUCTURE OF DISTRIBUTION CD	24

SECTION 1

INTRODUCTION

Air quality transport and diffusion models have been in use for several decades to estimate the impacts of pollutants released from point sources. Some of the simpler models are steady-state, Gaussian plume models that include little or no chemical transformation. During the model development phase and in subsequent upgrades, model performance is constantly evaluated.

In the past, the emphasis of the statistical evaluation comparisons has been on the "intended use." For instance, one of the uses for modeling results is to estimate the highest concentration values to be expected over a 5-year period, resulting from the operation of a proposed new power plant (Cox and Tikvart, 1990). Other statistical measures have also been employed to compare the concentration (or dose) values of "intended use," such as number of values within a factor of two, linear least-square fits to scatter plots of observed and predicted values, and normalized mean-squared errors of observed and predicted values (Fox, 1981). Implicit in such statistical comparisons is an assumption that the predicted and observed distributions of concentration values are from the same population, which may not be a well-founded assumption (Oreskes et al., 1994). Work is underway to develop a new generation of evaluation metrics that takes into account the statistical differences (in error distributions) between model predictions and observations. As a result, a shift in philosophy is occurring as to how models of environmental processes can be acceptably evaluated. Most models provide estimates of the first moment of conditions to be expected for each ensemble (e.g., average time-space variation of the meteorological conditions, average time-space variation of the surface-level concentration values). The key to the next-generation evaluation metrics is that they will no longer assume that the modeled and observed values come from the same statistical population of values. They will assume that they "share" certain fundamental properties, but are inherently different.

To fill a part of this void, the U.S. Environmental Protection Agency (EPA) has participated within the American Society for Testing and Materials (ASTM) to develop a consensus on a philosophical basis that could be used in developing new statistical methods for model evaluation. In doing this, a statistical procedure was drafted and documented within the Annex to Standard Guide D 6589 for Statistical Evaluation of Atmospheric Dispersion Model Performance (ASTM, 2000). This procedure implements the idea that the distributions of model predictions and observations "share" certain fundamental properties, but are inherently different. The procedure assesses the performance of transport and diffusion models to simulate the average "centerline" concentration values from a point source release.

The EPA has developed a FORTRAN program called "Design 2" that implements the

procedure described in the Annex to D 6589. This report describes the input files and formats of these input files for running Design 2. The report describes the information provided by the various output files created by Design 2. The report does not describe the actual statistical procedures of Standard Guide D 6589, as these are documented in detail within this Guide's Annex.

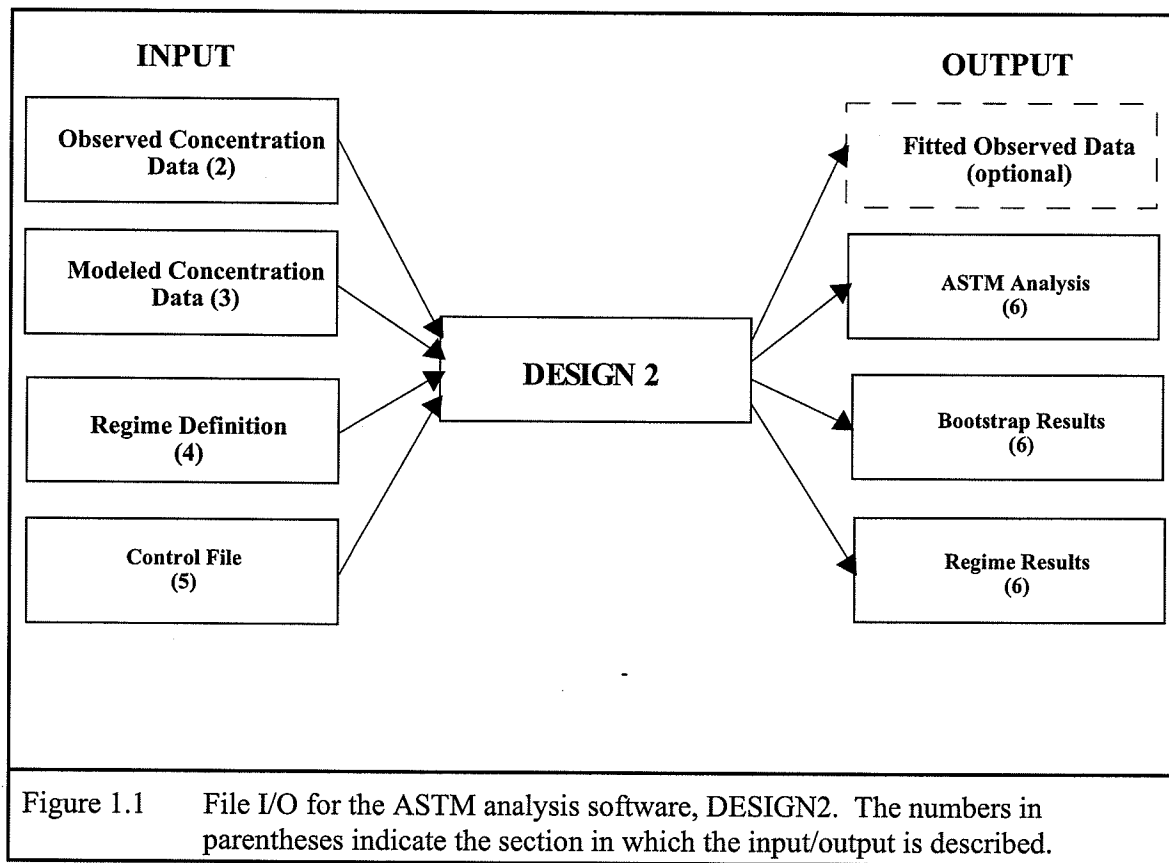
As noted in the ASTM standard, many practical problems must be resolved before the software can be run (see Section X1.4), including how to group the data into regimes. If we consider each observation period as an experiment and groups of monitors/receptors (nearly) equidistant downwind as an arc, then we can define an experiment-arc as a single experiment and single arc combination. An experiment-arc pair can then be included in a regime and the software can connect observed data to modeled data. Some regime definition strategies are mentioned in Section 4.

The purpose of this document is to provide guidance on the procedures to prepare the data and run the evaluation software. Figure 1.1 shows a flow chart of the process. The first input file required is the observed concentration data. The structure of this file is described in Section 2. The second file is the modeled concentration data. Since this software determines a best performing model, results from at least two models should appear in this file. The structure of this file is described in Section 3. The third file defines ranges of conditions, or regimes, based on one or more parameters that subdivide the results. In defining regimes, the behavior of the dispersion should be similar within each subdivision. A more complete discussion of regimes and the structure of this file is described in Section 4. For the software to function, a control file is required that has the names of the input and output file, file formats, and processing options. This file is short (less than 20 lines) and the structure of this file is described in Section 5. Once the software is run, several output files are produced: 1) the primary file with the information on the best performing model (the ASTM analysis), 2) a file of the bootstrap results, and 3) a file of results by regime. An optional output file is the fitted observed data. The analysis and structure of these files is discussed in Section 5. In Section 6, the results from the evaluation software are presented.

The software was coded in FORTRAN. Although a FORTRAN 95 compiler was used to create the executable, attempts were made to keep the code backward compatible with FORTRAN 77 standards. In several places, the user is required to provide a format for the data to be read by the program. These formats must follow FORTRAN conventions and MUST be enclosed in parentheses; for example, (1X, I5, I5, 3F12.0). The characters in the FORTRAN statement can be upper or lower case; the spaces after the commas are not necessary and are shown here only to provide better legibility.

The example used throughout this documentation is from the 1985 Indianapolis field experiment (Murray and Bowne, 1988). The field experiment was performed in a flat, urban to suburban environment. A tracer, SF₆, was injected into the buoyant gases exiting from an 86-meter point source. Data were collected for approximately 170 hours over a five-week period. Monitors were located in arcs from about 250 meters to 12 kilometers downwind.

Meteorological data were collected from a 94-meter and 10-meter tower in the urban environment, a 10-meter tower in a suburban location, and 10-meter tower in a rural location. National Weather Service meteorological data were from Indianapolis, IN (hourly weather observations) and Dayton, OH (upper air data). For the dispersion modeling, receptors were placed along a single line due east of the release point. To obtain centerline concentrations the wind directions were all modified to be blowing from the west (270°).



SECTION 2

OBSERVED CONCENTRATION DATA

The observed concentration data are required to be in a specific format for the analysis software to correctly read and process the data. The general structure is shown in Table 2.1. The first record in the file is Header 1 and contains a title that can be up to 40 characters long and only appears once. Note that this title must be enclosed in single quotation marks. Headers 2 and 3 appear before each arc of observations. These arc header records are followed by the x- and y-coordinates and observed concentration, for the number of receptors on the arc (the first field on Header 3). The coordinates and concentration must be specified as x-coordinate, y-coordinate and then concentration value. The FORTRAN format for this information must remain fixed for all values in the file. This FORTRAN format is defined by the user as part of the control file information (see Section 5). Currently, the number of observed concentration records per experiment-arc is limited to 250 records. An example of the observed concentration file is shown in Figure 2.1.

Header records 2 and 3 are read 'free' format, i.e., without a format statement. This means that a comma or at least one space (or a combination of the two, as seen in Figure 2.1) must separate each entry on each record. Since the date and time are the only character variables in these two records, they must be enclosed in single quotation marks.

The header 2 records provide general information on the field experiment period and arc combination. The experiment and arc (or traverse) pair identify an observation period and arc of data. The pair must be unique. Usually, the experiment numbers are unique (possibly a simple sequential numbering), but the arc numbers likely repeat from field experiment period to field experiment period. The important thing to remember is that these two parameters are used to connect the observed data with the modeled concentration data and the grouping into regimes, so the pair must be unique. The date, start and stop time, and nominal distance to the arc are not used in the program, but provide additional identifying information within the archive.

The header 3 records contain information on how to process this experiment-arc combination. The first three values define the total number of receptors on the arc and which receptors to use in the analysis. Receptors can be included or excluded from the analysis by varying the starting and ending point for the analysis.

The location of the receptors must be entered in the same coordinate system used in defining the receptor locations, namely Cartesian (x,y) or polar coordinates (r,θ). The type of coordinate system in use is defined by the value of IXYARC, which is the fourth parameter of

Header 3. If the receptor locations are entered as polar coordinates, then the program converts them to Cartesian coordinates internally. The program works with coordinates expressed in meters, so if the units are other than meters, a conversion factor from the user's units to meters must be applied. This factor is provided as the fifth parameter on header 3 records.

The sixth and seventh parameters identify the location of the source. If the x- and y-coordinates are entered as polar coordinates, then the source is likely to be located at an origin, i.e., (0,0). If the x- and y-coordinates are entered as Cartesian coordinates (e.g., UTM coordinates), then the source is likely to be located at something other than (0,0). The software translates the source and receptors to a (0,0)-based system internally to perform its analysis.

The tracer release rate is specified next and is used to normalize the fitted observed concentrations. The tracer height and the height of the receptors are not used in this version of the software, but may provide useful information possibly interpreting the results, and serve as a check on the source and receptor heights specified in the dispersion modeling.

The last parameter on this record converts the tracer concentration from user's units to micrograms per cubic meter ($\mu\text{g}/\text{m}^3$). Since this value is applied to every tracer concentration without regard to meteorology, careful consideration should be used as to the appropriate value to enter. For example, converting from parts per trillion to $\mu\text{g}/\text{m}^3$ requires atmospheric pressure and ambient temperature.

The header records are followed by the location and observed concentration for the number of receptors on the arc (defined by the first value on header type 3). At a minimum there are three values on these records. However, since the format is specified in the control file (see Section 5), any amount of information can be included and skipped with the format statement the user provides. For example, if the coordinate location is in polar coordinates, Cartesian coordinates can be included on the record but skipped with the appropriate format statement (as is done in Figure 2.1). During the fitting process, the concentrations are divided by the emission release rate (eighth parameter in Header 3).

TABLE 2.1 STRUCTURE OF THE OBSERVED CONCENTRATION DATA FILE.

Record Type	Data	Data Type	Description
Header 1	Title (up to 40 characters)	C	Appears only once, as the first record in the file, in single quotation marks
Header 2	Experiment number	I	Field experiment period
	Traverse number	I	Identifier for the arc
	Date	C	10 characters, in single quotes 'MM-DD-YY'
	Time (Start & Stop)	C	10 characters, in single quotes 'HHMM-HHMM'
	Nominal distance to arc	R	(not used)
Header 3	Number of receptors on the arc	I	Total number of receptors that will be read by the program
	Starting data point for integration	I	Receptors prior to this point are omitted from analysis
	Ending data point for integration	I	Receptors after this point are omitted from analysis
	How x- and y-coordinates are entered: polar or Cartesian (IXYARC)	I	<= 0, cartesian > 0, polar
	Conversion from user units to meters	R	Converts user units to meters, e.g., for kilometers to meters, enter 1000.
	X-coordinate of source	R	Coordinate in base system
	Y-coordinate of source	R	Coordinate in base system
	Tracer release rate	R	Release rate in grams/second
	Tracer release height	R	Height of release in meters (not used)
	Altitude of airborne traverse	R	Height of receptors in meters (not used)
	Constant multiplier to convert tracer concentration from user's units to $\mu\text{g}/\text{m}^3$	R	Enter 1.0 if no conversion is needed, otherwise enter the multiplier that will be applied to all observed concentrations
Observation	X-coordinate	R	If IXYARC <= 0, x is in meters If IXYARC > 0, x is in degrees
	Y-coordinate	R	If IXYARC <= 0, y is in meters If IXYARC > 0, y is radial distance to arc
	Observed Concentration	R	Concentration in user units

Data Types: C=character, I=integer, R=real

'IND SF6'									
1,	1,	' 09-16-85 '	' 1100-1200'		0.2				
1,	1,	1, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	24.71		.31	.13	.28				
1,	2,	' 09-16-85 '	' 1100-1200'		0.5				
8,	1,	8, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	5.10		.47	.04	.47				
	21.13		.47	.17	.44	155.0			
	34.63		.40	.23	.33	9.0			
	57.90		.45	.38	.24	.0			
	73.93		.43	.42	.12	.0			
	309.61		.50	-.39	.32	.0			
	325.45		.55	-.31	.45	18.0			
	347.88		.48	-.10	.47	143.0			
1,	3,	' 09-16-85 '	' 1100-1200'		0.7				
4,	1,	4, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	38.25		.66	.41	.52	.0			
	332.79		.71	-.32	.63	38.0			
	338.05		.68	-.25	.63	25.0			
	356.90		.72	-.04	.72	84.0			
1,	4,	' 09-16-85 '	' 1100-1200'		1.0				
6,	1,	6, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	250.85		1.43	-1.35	-.47	.0			
	290.40		1.03	-.97	.36	.0			
	298.74		.94	-.82	.45	.0			
	323.61		1.01	-.60	.81	6.0			
	335.41		1.01	-.42	.92	116.0			
	344.65		1.02	-.27	.98	206.0			
1,	5,	' 09-16-85 '	' 1100-1200'		1.5				
11,	1,	11, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	6.07		1.41	.15	1.40	225.0			
	15.99		1.48	.41	1.42	59.0			
	257.66		1.50	-1.46	-.32	.0			
	263.77		1.47	-1.46	-.16	.0			
	269.59		1.38	-1.38	-.01	9.0			
	280.49		1.81	-1.78	.33	.0			
	287.25		1.52	-1.45	.45	.0			
	293.21		1.55	-1.42	.61	.0			
	302.81		1.77	-1.49	.96	.0			
	341.93		1.47	-.46	1.40	79.0			
	358.25		1.47	-.05	1.47	444.0			
1,	6,	' 09-16-85 '	' 1100-1200'		2.0				
17,	1,	17, 1,	1000.000,	0.0,	0.0,	4.94000,	83.80,	1.50,	1.000
	10.43		1.97	.36	1.94	55.0			
	27.36		1.93	.88	1.71	.0			
	44.34		1.92	1.34	1.37	.0			
	51.53		1.95	1.52	1.21	.0			
	68.32		1.92	1.79	.71	.0			
	250.99		2.06	-1.94	-.67	.0			
	257.43		1.97	-1.93	-.43	.0			
	268.08		1.79	-1.79	-.06	.0			
	274.44		2.20	-2.19	.17	.0			
	282.63		1.97	-1.92	.43	.0			
	290.17		2.03	-1.91	.70	.0			
	297.51		2.01	-1.79	.93	.0			
	305.40		2.07	-1.69	1.20	.0			
	313.61		2.04	-1.48	1.41	.0			
	323.38		2.13	-1.27	1.71	6.0			
	332.62		1.95	-.90	1.73	15.0			
	340.45		1.96	-.66	1.85	40.0			

Figure 2.1. Sample of the observed concentration file.

SECTION 3

MODELED CONCENTRATION DATA

The second component required to perform the analysis is a file of modeled concentration data. This file contains the concentration estimates from two or more dispersion models. Results from as many model runs can be included (within DOS or compiler limitations) in this file since the user must specify in the control file the number of model runs to include in the analysis and the format to read the data. However, only results from up to 15 models can be included in a single analysis. The structure of the file is shown in Table 3.1 and an example is shown in Figure 3.1.

The format of these data is contained in the control file (described in Section 5). Specifying the format in this manner allows the user to identify which model estimates to include in the analysis.

It should be mentioned that the units of the modeled concentration estimates **MUST** match those in the fitted observed data, otherwise the results will be completely unreliable. It may be necessary to develop an external program that converts the modeled concentrations to the units in the observed concentration file, as was done for the example used in this documentation.

TABLE 3.1 STRUCTURE OF MODELED CONCENTRATION DATA FILE.

Record	Description
1	Header record; can be any descriptive language; the information in this record is not used in the program
2	Header record; can be any descriptive language; the information in this record is not used in the program
3 ... # of experiment-arc periods	Modeled concentration data; multiple columns of data; repeated for as many arc-experiment periods to be included in the analysis Required data on each record for N model runs (N is specified in the control file): Experiment period, Arc number, Concentration 1, ..., Concentration N

(3x,2i5,f11.3,2x,i6.6,i4.4,5f12.5)								
EXP	TRAV	X	DATETIME	ADMS	AER02161	HPDM	ISCST3	AMODNoObsZi
1	1	0.200	8509161100	13.52654	15.88650	20.58027	0.00000	22.98427
1	2	0.500	8509161100	124.19214	62.27761	120.69160	11.44278	67.62427
1	3	0.700	8509161100	129.02158	62.43041	115.09613	51.72357	65.56213
1	4	1.000	8509161100	106.83546	51.13265	88.37265	76.59892	51.24015
1	5	1.500	8509161100	72.80358	35.57238	55.99589	68.48630	32.48832
1	6	2.000	8509161100	52.18434	27.28770	40.92309	57.06482	22.16314
1	7	3.000	8509161100	31.09959	18.67307	27.57795	42.60660	13.06510
1	8	4.000	8509161100	21.46810	14.77599	21.56590	34.89255	9.14368
1	9	6.000	8509161100	13.23311	11.21865	16.25815	26.69948	5.99775
1	10	8.000	8509161100	9.56996	9.23732	13.34078	22.29126	4.77061
1	11	10.000	8509161100	7.39425	7.90674	11.56248	19.47426	4.03602
1	12	12.000	8509161100	5.92507	6.94058	10.34749	17.48786	3.52151
2	1	0.200	8509161200	9.67782	17.32178	16.85009	0.00000	19.70494
2	2	0.500	8509161200	116.65285	67.36076	117.75447	10.78590	67.53592
2	3	0.700	8509161200	122.54763	67.21519	112.84499	48.39931	66.77747
2	4	1.000	8509161200	101.43091	53.96894	85.90490	70.07381	52.49763
2	5	1.500	8509161200	68.40762	35.32141	52.63741	54.24477	33.22327
2	6	2.000	8509161200	48.55746	24.98616	34.97790	41.89233	22.42547
2	7	3.000	8509161200	28.52670	15.65525	20.92225	30.92515	12.81057
2	8	4.000	8509161200	19.32401	11.41748	15.39928	25.32321	8.82312
2	9	6.000	8509161200	11.43619	8.05074	11.10693	19.37537	5.54923
2	10	8.000	8509161200	8.15481	6.53249	9.14303	16.17562	4.33742
2	11	10.000	8509161200	6.31754	5.56385	7.83048	14.13122	3.65312
2	12	12.000	8509161200	5.09892	4.87275	6.94412	12.68949	3.18173
3	1	0.200	8509161300	6.93544	16.03329	11.07608	0.00000	17.17133
3	2	0.500	8509161300	110.24163	69.93465	102.03912	11.46303	68.71230
3	3	0.700	8509161300	120.36591	70.52279	103.26080	48.85902	69.20486
3	4	1.000	8509161300	99.47187	56.41741	81.77794	68.63754	54.65300
3	5	1.500	8509161300	66.36179	36.37708	51.53038	50.91323	34.47306
3	6	2.000	8509161300	46.57782	24.99461	34.29897	37.08468	23.10219
3	7	3.000	8509161300	26.95085	14.90788	19.38106	26.71908	12.85674
3	8	4.000	8509161300	18.02947	10.54413	13.86609	21.87267	8.71999
3	9	6.000	8509161300	10.29495	7.00644	9.48170	16.73271	5.30995
3	10	8.000	8509161300	7.16038	5.59702	7.80348	13.96850	4.06644
3	11	10.000	8509161300	5.51101	4.74401	6.67414	12.20238	3.40525
3	12	12.000	8509161300	4.45420	4.14563	5.88584	10.95750	2.95946
4	1	0.200	8509161400	6.02123	0.01478	3.08309	0.00000	0.01169
4	2	0.500	8509161400	109.91783	50.51303	76.51933	21.24328	48.07773
4	3	0.700	8509161400	122.56912	89.26309	92.48806	64.44592	86.99314
4	4	1.000	8509161400	102.79445	84.29244	82.91390	75.67673	83.30546
4	5	1.500	8509161400	67.80977	57.40505	56.95367	50.77602	56.88217
4	6	2.000	8509161400	47.12925	39.04949	39.30663	34.06233	38.64830
4	7	3.000	8509161400	26.90252	20.83178	21.43528	23.24096	20.55332
4	8	4.000	8509161400	17.84781	12.90603	14.23599	18.99567	12.70320
4	9	6.000	8509161400	9.99287	6.44580	8.91191	14.52074	6.32857
4	10	8.000	8509161400	6.80607	3.97508	6.91433	12.11810	3.90839
4	11	10.000	8509161400	5.18173	2.83168	5.94636	10.58417	2.79318
4	12	12.000	8509161400	4.17275	2.22595	5.25265	9.50299	2.20017

Figure 3.1. Example of modeled concentration file

The first two records can contain any information the user wants to include since neither record is used by the analysis software. In this example, the FORTRAN format of the data appears on the first record. The second record identifies the fields. In this example, there are two model runs with the AERMOD dispersion model: the first with the column header AER02161 and a second with the column header AMODNoObsZi. The former run utilized observed mixing heights, and the latter allowed AERMET to calculate the mixing heights.

SECTION 4

REGIMES

The third file required to perform the ASTM analysis is a file of 'regime' definitions. This file groups the experiment-arc combinations into categories that the user feels are appropriate to answer the question(s) under investigation. Comparing separately averaged observed and modeled concentrations within grouped data "provides an empirical estimate of the combined deterministic error associated with input uncertainty and formulation errors" (ASTM, 2000). An extended discussion of establishing the regimes is provided in the ASTM standard guide (ASTM, 2000).

Regimes can be defined based on almost any property, such as atmospheric stability, type of air mass, or wind speed. The definitions of the regimes appearing in the example in this document were based on z_i/L , where z_i is the mixing height and L is the Monin-Obukhov length. Two additional restrictions were placed on the data to insure there were sufficient data to perform the analysis. The first was that a minimum of five nonzero observed concentrations must appear on an arc to be included in a regime and that there must be a minimum of five experiment-arcs per regime. Regimes, including the one presented here, are not hard and fast definitions, and in fact, may require modification if the initial results are ambiguous or the meteorology changes. Additionally, which model's meteorology (or other model-specific data) to use to define the regimes must be considered when developing the regimes.

The structure of the regime definition file is shown in Table 4.1. There are three record types, with the first type defining the number of regimes (and appearing only once, at the beginning of the file), and the second type defining the number of experiment-arc combinations in a regime. The third record type consists of the experiment-arc pairs within the regime. All required data are read as free format integers.

An example of the regime definition file is shown in Figure 4.1. Note that in the example, additional information follows the required data on each record. Only the first field is read on record types 1 and 2, so any descriptive language can follow these values. For example, the criteria for the data grouping can appear after the first field. In record type three, only the first two fields are required and are read free format. Additional information can appear after these two fields that may be useful in reviewing and understanding the results. In this particular example, z_i/L is used as the criterion to group the data. The first regime is defined for $z_i/L < -50.0$. The values that follow the experiment and arc number on each of the type 3 records is the number of nonzero observed concentrations.

TABLE 4.1 STRUCTURE OF THE REGIME DEFINITION FILE

Record Type	Description
1	Number of regimes to use; appears only once as the first record
2	Number of experiments-arc combinations in the regime; appears before each grouping of experiment-arcs
3	Experiment period and arc to include in this regime

```

29 Number of Regimes
12 Number of Exps in this regime Cases where Zi/L is less than: -50.0
  2   3   6
  3   3   7
19   3   8
20   3   7
21   3   6
28   3   6
32   3   9
33   3   8
34   3  13
113  3
114  3  10
116  3   5
12 Number of Exps in this regime
  2   4   9
  3   4   8
19   4   6
20   4   6
21   4   7
28   4   5
32   4   9
33   4   7
34   4  13
113  4
114  4  15
116  4   8
.     .   .
.     .   .
.     .   .
9 Number of Exps in this regime Cases where Zi/L is greater than: -50.0 and less
than: -25.0
22   3   9
23   3   9
24   3   7
25   3   8
29   3   7
30   3   8
31   3  10
35   3   5
55   3   5
14 Number of Exps in this regime
22   4   8
.     .   .
.     .   .
.     .   .

```

Figure 4.1. Example of the regime definition file. The required fields and information are shown in bold and the information that is not required is shown in a lighter font.

SECTION 5

CONTROL FILE

Once the data files are in place, the remaining file required to run the ASTM evaluation software is the control file. This file contains file names, data formats, and processing options. Table 5.1 shows the structure of the control file and Figure 5.1 shows an example. The output from the analysis software will be presented in Section 6.

The first six records in the control file pertain to the fitting of the observed concentration data. The fourth record is needed only if the results of the fit are to be retained in a file. Otherwise, the data are retained in memory for the remainder of the analysis and lost when the program terminates. The fifth record (IPHIY) indicates how σ_y is to be retained, either in user units or in degrees. The sixth record defines the minimum number of nonzero concentration values required to attempt a Gaussian fit of the observed data.

The remaining controls are relevant to the analysis of the modeled data and the actual evaluation analysis. Records 7-9 apply to the sampling scheme (NPAIR) and method (NWIDE), and the number of bootstrap samples (NBOOT). NPAIR defines whether to sample the data as pairs or individually; NWIDE defines whether to sample modeled and observed concentrations simultaneously or independently.

Record 10 defines the number of models or models runs (NMODEL) that are to be analyzed with this run. The analysis can be performed with up to 15 models in any single run. Note too that different runs of the same model or versions of a model can be included in an analysis. This record is followed by the names of the NMODEL models. The names are only used to identify the models in the output files. It is recommended to keep the identifying names to 10 characters or less. If the name is longer than 10 characters, only the first 10 characters will be used. The order of the names must match the order the modeling results are read in the modeled concentration file (Section 3).

The record following the model names is the random number seed and should be at least five digits. This value is used during the bootstrapping process. The seed value is followed by the filtering option. It is an integer value greater than or equal to zero and controls the maximum number of near-centerline concentrations to be considered that satisfy $-0.67 < y/\sigma_y < 0.67$. Entering a value of 0 selects all near-centerline concentrations that satisfy this criterion.

The last six records (shown in Table 5.1 as records 13+NMODEL through 18+NMODEL) in the control file are input and output file names and the format of the modeled concentration file. The names should follow the standard personal computer (PC) naming

conventions. The FORTRAN code and executable version accompanying this documentation accommodates filenames as long as 70 characters in length. The format specified on the fifth record from the end can be as long as 30 characters in length, must follow FORTRAN format statement rules for the data to be read correctly, and must include the initial and closing parentheses.

TABLE 5.1 STRUCTURE OF THE INPUT CONTROL FILE FOR THE ASTM EVALUATION SOFTWARE.

Record	Control*	Description
1	filename	File where arcs of observed data are stored
2	FRMYC	Format to read the location (x- and y-coordinates) and observed concentration (can be up to 30 characters in length, including parentheses).
3	ISPLUS	Option to save results from the PLTFIT routines in a file: = 0, do not save results to file ≠ 0, save results to file
4	output filename	File with fitted observed data - an optional record based on the value given on the 3 rd record
5	IPHIY	σ_y output units: = 0, list output σ_y in user units, ≠ 0, list σ_y in degrees
6	MINNOK	Minimum number of values required to attempt Gaussian fit
7	NPAIR	Sampling scheme: = 2, sample by pairs ≠ 2, do not sample by pairs
8	NWIDE	Sampling method = 2, sample simultaneously observation & modeled values ≠ 2, do not sample simultaneously
9	NBOOT	Number of bootstrap samples
10	NMODEL	Number of models for which values will be provided
11	model name	First model name
:	:	:
10+NMODEL	model name	Last model name
11+NMODEL	ISEED	Random number seed value (make sure it is 5 or more digits)
12+NMODEL	NFILTER	Filtering option (if zero, then no limit; otherwise this is the max number of values to be selected)
13+NMODEL	input filename	File of modeled centerline concentration values for each model
14+NMODEL	data format	Format of the data of modeled centerline concentration values
15+NMODEL	input filename	File defining which experiment numbers and arcs are in each Regime
16+NMODEL	output filename	Name of output file for listing results - the primary listing for evaluating model results with the ASTM standard
17+NMODEL	output filename	File with bootstrap results
18+NMODEL	output filename	File with regime results

* Controls in capital letters show the variable name used in the source code, otherwise a generic 'name' is used.

```
INDARCS.DAT
(1x,F8.0,F8.0,16x,F8.0)
1
INDSPLUS.DAT
2
3
2
2
500
5
ADMS
AMOD02161
HPDM
ISCST3
AMODNoObzi
12345
1
INDMODEL.DAT
(3x,2i5,13x,10x,5f12.0)
INDREGM.DAT
INDASTM.OUT
INDBOOT.PLT
INDREGIM.OUT
```

Figure 5.1 Sample control file.

SECTION 6

RESULTS

In the earlier sections, portions of the various input files were shown. In this section, the results from the analysis software are presented based on the example files presented in the previous sections. Due to the size of the files, not all sections of the output files will be shown. A brief description of the general layout of the files will be provided.

The output filenames are specified by the user, as shown in Table 5.1 and Figure 5.1. The last two files in the control file are results from the bootstrap method and information by regime. The third file from the end (identified as 16+NMODEL in Table 5.1, INDASTM.OUT in Figure 6.1) contains the listing of results used to apply the ASTM standard. The first part of this file has the following structure:

- 1) Run summary of the information provided in the control file;
- 2) Results for each regime, including the number of nonzero values for each experimental period and the number of centerline values that satisfy the following $-0.67 < y/\sigma_y < 0.67$, where y is the receptor distance from centerline and σ_y is the standard deviation of the distances for all nonzero concentration estimates;
- 3) Average and standard deviation of C_i , and the geometric average and standard deviation of C_i/C_{avg} , where the near-centerline concentrations and averages use only the values resulting from the filtering process (i.e., limit on number of near-centerline concentrations).

Items 2) and 3) are repeated for all regimes. Figure 6.1 shows an example of these initial records in the output file.

ASTM D22.11 Boot.for Draft Z6849Z VERSION: 063002

SETUP INFORMATION:

Iseed (random number seed value): 12345
 Npair (0=individual,2=pair sampling): 2
 Nwide (0=individual,2=concurrent): 2
 Nboots (number of bootstrap samples): 500
 Nmodel (number of models): 5
 Model 1 ADMS
 Model 2 AMOD02161
 Model 3 HPDM
 Model 4 ISCST3
 Model 5 AMODNoObZi
 Iphiy (0=m, 1=degrees): 1
 Nfilter (Limit on number selected): 1
 Model file: indmodel.dat
 Output file: indASTM.out
 Regime Definition File: indregm.dat

Results for Group: 1

N	Exp	Arc	NumValues	NumCenterline	NumModels
1	2	3	13	1	5
2	3	3	14	1	5
3	19	3	17	1	5
4	20	3	19	1	5
5	21	3	18	1	5
6	28	3	15	1	5
7	32	3	18	1	5
8	33	3	18	1	5
9	34	3	18	1	5
10	113	3	10	1	5
11	114	3	10	1	5
12	116	3	10	1	5

Total number of values: 180
 Computed Group Sy: 19.10 (deg)

	Num	Avg Ci	Std Ci	Num >0	GeoAvg Ci/Cavg	GeoStd Ci/Cavg
--OBS	12	5.86E+01	4.49E+01	11	6.30E-01	3.31E+00
ADMS	12	1.28E+02	4.63E+01	12	7.41E-01	3.37E+00
AMOD02161	12	7.41E+01	1.77E+01	12	9.63E-01	1.34E+00
HPDM	12	1.20E+02	3.64E+01	12	9.26E-01	1.58E+00
ISCST3	12	7.17E+01	3.94E+01	10	1.14E+00	1.37E+00
AMODNoObZi	12	6.09E+01	3.73E+01	12	4.17E-01	9.58E+00

Figure 6.1. The initial records in the output file used to determine the best performing model (associated with record NMODEL+16 in the control file).

These statistics by regime are followed by the statistical measures over all regimes for each model and maximum observed value. Currently, there are 12 statistical measures calculated in the software: fractional bias, absolute fractional bias, normalized mean squared error (NMSE), mean squared error (MSE), slope, intercept, r^2 , (unsystematic MSE)/MSE, (systematic MSE)/MSE, Willmott-d, average observed value, and average modeled value. For each statistical measure, an abbreviated description is provided of the distribution of values generated during the bootstrap: average, standard deviation, minimum, low hinge (25-th percentile), median, high hinge (75-th percentile), and maximum value. It is these statistics that are used to determine the "best performing model." An example of the summary information provided for each statistical measure is shown in Figure 6.2.

For each statistical measure, the "best performing model" is indicated by an asterisk or an exclamation point after the model name (and before the column labeled Avg). An asterisk is used for those measures for which the smallest (absolute) average value identifies the best performing model. The exclamation point is used for those measures for which a value closest to 1.0 is appropriate, such as the correlation coefficient and the slope. Note that the first three and the last three statistical measures are shown in the figure; there are another six measures not shown.

Summary Over All Groups (No Inverse Variance Weighting Employed)									
Fractional Bias Results									
Model	Avg	Std	Nboots	Min	HL	Med	HH	Max	
MAX	0.11905	0.02346	500	0.06918	0.10212	0.11892	0.13289	0.20091	
ADMS	0.00008	0.06137	500	-0.18474	-0.04431	-0.00080	0.04321	0.22366	*
AMOD02161	-0.29222	0.06333	500	-0.52664	-0.33212	-0.29712	-0.25111	-0.09162	
HPDM	-0.03237	0.06574	500	-0.22245	-0.08129	-0.03377	0.01350	0.17498	
ISCST3	0.03411	0.05951	500	-0.17623	-0.00250	0.03322	0.07034	0.21014	
AMODNoObZi	-0.22308	0.06185	500	-0.40081	-0.26444	-0.22674	-0.17864	-0.02809	
t-value compared to the base model: ADMS									
2 AMOD02161	5.7244								
3 HPDM	0.8565								
4 ISCST3	0.8028								
5 AMODNoObZi	5.5705								
Degrees of freedom (DF): 28									
(at the 90% confidence level a value > 1.701 indicates that there is a significant difference in the metrics between the base model and the model being tested)									
Absolute Fractional Bias Results									
Model	Avg	Std	Nboots	Min	HL	Med	HH	Max	
MAX	0.11905	0.02346	500	0.06918	0.10212	0.11892	0.13289	0.20091	
ADMS	0.55608	0.04422	500	0.43668	0.52596	0.55828	0.58757	0.68484	
AMOD02161	0.60549	0.04837	500	0.46769	0.57155	0.60523	0.63789	0.74218	
HPDM	0.46838	0.04687	500	0.30258	0.43679	0.47103	0.50124	0.62682	*
ISCST3	0.48538	0.05049	500	0.32799	0.45442	0.48618	0.51749	0.66683	
AMODNoObZi	0.54344	0.04826	500	0.40345	0.51097	0.54504	0.57555	0.67938	
t-value compared to the base model: HPDM									
1 ADMS	2.0452								
2 AMOD02161	3.0677								
4 ISCST3	0.3725								
5 AMODNoObZi	1.6021								
Degrees of freedom (DF): 28									
(at the 90% confidence level a value > 1.701 indicates that there is a significant difference in the metrics between the base model and the model being tested)									
Normalized Mean Squared Error Results									
Model	Avg	Std	Nboots	Min	HL	Med	HH	Max	
MAX	0.02203	0.00956	500	0.00478	0.01491	0.02013	0.02714	0.06065	
ADMS	0.51363	0.06581	500	0.33369	0.46615	0.51186	0.55525	0.72839	
AMOD02161	0.45289	0.09301	500	0.24901	0.38521	0.44552	0.51077	0.84992	
HPDM	0.34060	0.07401	500	0.10480	0.29104	0.33899	0.38767	0.63089	*
ISCST3	0.42333	0.11733	500	0.18697	0.34874	0.40906	0.48065	1.04803	
AMODNoObZi	0.42698	0.09067	500	0.22267	0.36160	0.42129	0.48415	0.79317	
t-value compared to the base model: HPDM									
1 ADMS	2.7531								
2 AMOD02161	1.4852								
4 ISCST3	0.7391								
5 AMODNoObZi	1.0565								
Degrees of freedom (DF): 28									
(at the 90% confidence level a value > 1.701 indicates that there is a significant difference in the metrics between the base model and the model being tested)									
Willmott - d Results									
Model	Avg	Std	Nboots	Min	HL	Med	HH	Max	
MAX	0.97037	0.01428	500	0.91508	0.96186	0.97311	0.98161	0.99403	
ADMS	0.59000	0.05730	500	0.41442	0.55072	0.59026	0.62879	0.76611	
AMOD02161	0.61483	0.06993	500	0.42199	0.56497	0.61813	0.66311	0.80535	
HPDM	0.64662	0.07320	500	0.46643	0.59056	0.64853	0.69900	0.89654	!
ISCST3	0.59026	0.08255	500	0.33910	0.53556	0.59158	0.65071	0.78874	
AMODNoObZi	0.60632	0.07067	500	0.42730	0.55701	0.60583	0.65867	0.82291	
t-value compared to the base model: HPDM									
1 ADMS	1.1011								
2 AMOD02161	0.5605								
4 ISCST3	0.7755								
5 AMODNoObZi	0.6283								
Degrees of freedom (DF): 28									
(at the 90% confidence level a value > 1.701 indicates that there is a significant difference in the metrics between the base model and the model being tested)									
Average of Observed Values Results									
Model	Avg	Std	Nboots	Min	HL	Med	HH	Max	
MAX	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	
ADMS	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	
AMOD02161	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	
HPDM	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	
ISCST3	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	
AMODNoObZi	49.65137	2.83214	500	42.61565	47.59573	49.72298	51.62337	59.85370	

Figure 6.2. A portion of the output file with the statistical measures for the ASTM analysis. The “best performing model” is indicated by an asterisk or exclamation point for each statistical measure.

To assess whether or not there is a significant difference in the statistical measure between the "best performing model" and the other models, the bootstrap results for the statistical measure are used to compute the t-value(s). The null hypothesis that the average of the model bootstrap differences is greater than 0 can be tested for a given confidence level (ASTM, 2000). The t-values are shown immediately below the statistical measures in Figure 6.2. Examining the Normalized Mean Square Error in Figure 6.2, HPDM is the "best performing model," but there is no statistical difference at the 90% confidence level between HPDM and both versions of AERMOD and ISCST3. However, there is a significant difference between HPDM and ADMS.

The other two output files are a file of bootstrap results and results by regime. Each contains different information that may be useful in interpreting the results. The data used to compute the t-value(s) come from the bootstrap file.

The file of bootstrap results is divided into five parts:

- 1) Run summary of the information provided in the control file;
- 2) Averages and standard deviations of observed and modeled concentrations by regime;
- 3) Averages and standard deviations of the fractional bias and absolute fractional bias for each regime. The definitions of fractional bias can be found in the discussion of the ASTM standard (ASTM, 2000);
- 4) Averages computed for each boot and regime;
- 5) Comparison statistics for each boot; for each statistical measure.

In this example, there are nearly 21,000 records in the output file. Due to the length of the initial records and the number of records in the file, a sample is not shown here.

Figure 6.3 shows a sample of the information by regime. The standard deviation of the concentration values along the arc is listed first. Then for each experiment-arc combination an analysis is provided of the observed concentration values satisfying the criterion of being within $-0.67 < y/\sigma_y < 0.67$. C_i is a concentration value for this arc that satisfies the requirements of NFILTER and C_{avg} is the average of all N values with zero values excluded for this regime. The asterisk next to a C_i value indicates this value is the maximum observed concentration found anywhere along the arc. There can be multiple occurrences for an experiment-arc pair depending on the value of NFILTER and the number values that are within $\pm 0.67 y/\sigma_y$. Some additional statistics appear below this grouping and are based on the nonzero values. These statistics are followed by similar results where NFILTER was applied in the modeling. In this example, NFILTER was set to one, so only one near-centerline concentration is selected for analysis.

```

Process Regime:          1
Sy(deg): 19.10201
  i of N Exp Arc      Ci      Cavg      Ci/Cavg
1  27  2  3      48.5830      58.3952      0.8320
2  27  2  3      92.7125*      58.3952      1.5877
3  27  2  3      35.4251      58.3952      0.6066
4  27  3  3      78.5425*      58.3952      1.3450
5  27 19  3      34.5494*      58.3952      0.5916
6  27 19  3      26.8240      58.3952      0.4594
7  27 19  3      30.2575      58.3952      0.5182
8  27 20  3      53.8627      58.3952      0.9224
9  27 20  3     108.7983*      58.3952      1.8631
10 27 20  3      72.9614      58.3952      1.2494
11 27 21  3     148.1799*      58.3952      2.5375
12 27 21  3     134.4754      58.3952      2.3029
13 27 28  3          0.0000      58.3952      0.0000
14 27 28  3          0.0000      58.3952      0.0000
15 27 28  3      30.5376*      58.3952      0.5229
16 27 32  3      60.6452*      58.3952      1.0385
17 27 32  3      56.7742      58.3952      0.9722
18 27 33  3      72.4731      58.3952      1.2411
19 27 33  3     109.6774*      58.3952      1.8782
20 27 33  3      68.3871      58.3952      1.1711
21 27 34  3      21.7204      58.3952      0.3720
22 27 34  3      13.7634      58.3952      0.2357
23 27 113 3          5.5794      58.3952      0.0955
24 27 113 3          5.7940      58.3952      0.0992
25 27 114 3          3.2189      58.3952      0.0551
26 27 116 3      81.3305*      58.3952      1.3928
27 27 116 3      64.8069      58.3952      1.1098

  Avg      Std      GeoAvg      GeoAvg
  Ci      Ci      Ci/Cavg      Ci/Cavg
54.0696  40.2394      0.7105      2.6809

```

Results from Subroutine Filter

```

  i of N Exp Arc      Ci      Cavg      Ci/Cavg
1  12  2  3      92.7125*      63.9209      1.4504
2  12  3  3      78.5425*      63.9209      1.2287
3  12 19  3      26.8240      63.9209      0.4196
4  12 20  3     108.7983*      63.9209      1.7021
5  12 21  3     134.4754      63.9209      2.1038
6  12 28  3          0.0000      63.9209      0.0000
7  12 32  3      56.7742      63.9209      0.8882
8  12 33  3     109.6774*      63.9209      1.7158
9  12 34  3      21.7204      63.9209      0.3398
10 12 113 3          5.5794      63.9209      0.0873
11 12 114 3          3.2189      63.9209      0.0504
12 12 116 3      64.8069      63.9209      1.0139

```

```

  Avg      Std      GeoAvg      GeoAvg
  Ci      Ci      Ci/Cavg      Ci/Cavg
58.5942  44.8805      0.6298      3.3111

```

```

Process Regime:          2
Sy(deg): 19.13367
  i of N Exp Arc      Ci      Cavg      Ci/Cavg
1  30  2  4      66.1943*      54.5058      1.2144
2  30  2  4      41.9028      54.5058      0.7688
.      .      .      .
.      .      .      .

```

Figure 6.3. Sample of the output by regime.

SECTION 7

REFERENCES

ASTM, 2000: Standard Guide for Statistical Evaluation of Atmospheric Dispersion Model Performance. D 6589-00. American Society for Testing and Materials, West Conshohocken, PA.

Cox, W. M. and J. A. Tikvart, 1990: A Statistical Procedure for Determining the Best Performing Air Quality Model, *Atmospheric Environment*, **24A**, 2387-2395.

Fox, D. G., 1981: Judging Air Quality Model Performance, *Bulletin of the American Meteorological Society*, **62**, 599-609.

Murray, D. R. and N. E. Bowne, 1988: Urban Power Plant Plume Studies. EPRI Report No. EA-5468, Research Project 2736-1, Electric Power Research Institute, Palo Alto, CA.

Oreskes, N., K. Shrader-Frechette and K. Belitz (1994): Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263: 641-646.

APPENDIX

DIRECTORY STRUCTURE OF DISTRIBUTION CD

ASTM Root Directory

Indianapolis

ASTM Evaluation

Dispersion Modeling

ADMS

AERMOD

AERMOD_NoObZi

HPDM

ISCST3

Meteorology

ADMS

AERMOD

AERMOD_NoObZi

HPDM

ISCST3

Kincaid

ASTM Evaluation

Dispersion Modeling

ADMS

AERMOD

AERMOD_Turb

HPDM

ISCST3

Meteorology

ADMS

AERMOD

AERMOD_Turb

HPDM

ISCST3

MixHts

PrairieGrass

ASTM Evaluation

Dispersion Modeling

ADMS

AERMOD

HPDM

ISCST3

Meteorology

ADMS

AERMOD

HPDM

ISCST3

MixHts

Documentation

Software

TECHNICAL REPORT DATA <i>(Please read Instructions on reverse before completing)</i>		
1. REPORT NO. EPA-454/R-03-006	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE A Program for the Statistical Evaluation Of Point Source Dispersion Models Using ASTM D 6589	5. REPORT DATE August 2003	
	6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S)	8. PERFORMING ORGANIZATION REPORT NO.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Pacific Environmental Services, Inc, a MACTEC Company Research Triangle Park, NC	10. PROGRAM ELEMENT NO.	
	11. CONTRACT/GRANT NO.	
12. SPONSORING AGENCY NAME AND ADDRESS Office of Air Quality Planning and Standards Emissions, Monitoring, and Analysis Division U.S. Environmental Protection Agency Research Triangle Park, NC 27711	13. TYPE OF REPORT AND PERIOD COVERED Technical Report	
	14. SPONSORING AGENCY CODE	
15. SUPPLEMENTARY NOTES		
16. ABSTRACT The ASTM published a Standard Guide D 6589 for Statistical Evaluation of Atmospheric Dispersion Model Performance, which provides a procedure for assessing performance of atmospheric transport and diffusion models to predict the "average centerline" concentration values of a chemical released from a point source. The EPA has developed a FORTRAN program entitled "Design 2" that implements this procedure. This report describes the input files and formats of these input files for running Design 2. The report describes the information provided by the various output files created by Design 2. The report does not describe the actual statistical procedures of Standard Guide D 6589, as these are documented in detail within this Guide's Annex.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Meteorological Data Air Dispersion Models	Statistical Model Evaluation	
18. DISTRIBUTION STATEMENT Release Unlimited	19. SECURITY CLASS (<i>Report</i>) Unclassified	21. NO. OF PAGES 34
	20. SECURITY CLASS (<i>Page</i>) Unclassified	22. PRICE

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Planning and Standards
Research Triangle Park, NC 27711

Publication No. EPA-454/R-03-006
August 2003

WASH DC - HQ
EPA-454/R-03-006