Recovery and Use of Existing Atmospheric Transport and Dispersion (ATD) and Associated Meteorological Data

Joseph C. Chang
Homeland Security Institute
Arlington, VA

July 19, 2005

OFCM Special Session
9th George Mason University Annual Conference on
Atmospheric Transport and Dispersion Modeling
Fairfax, Virginia

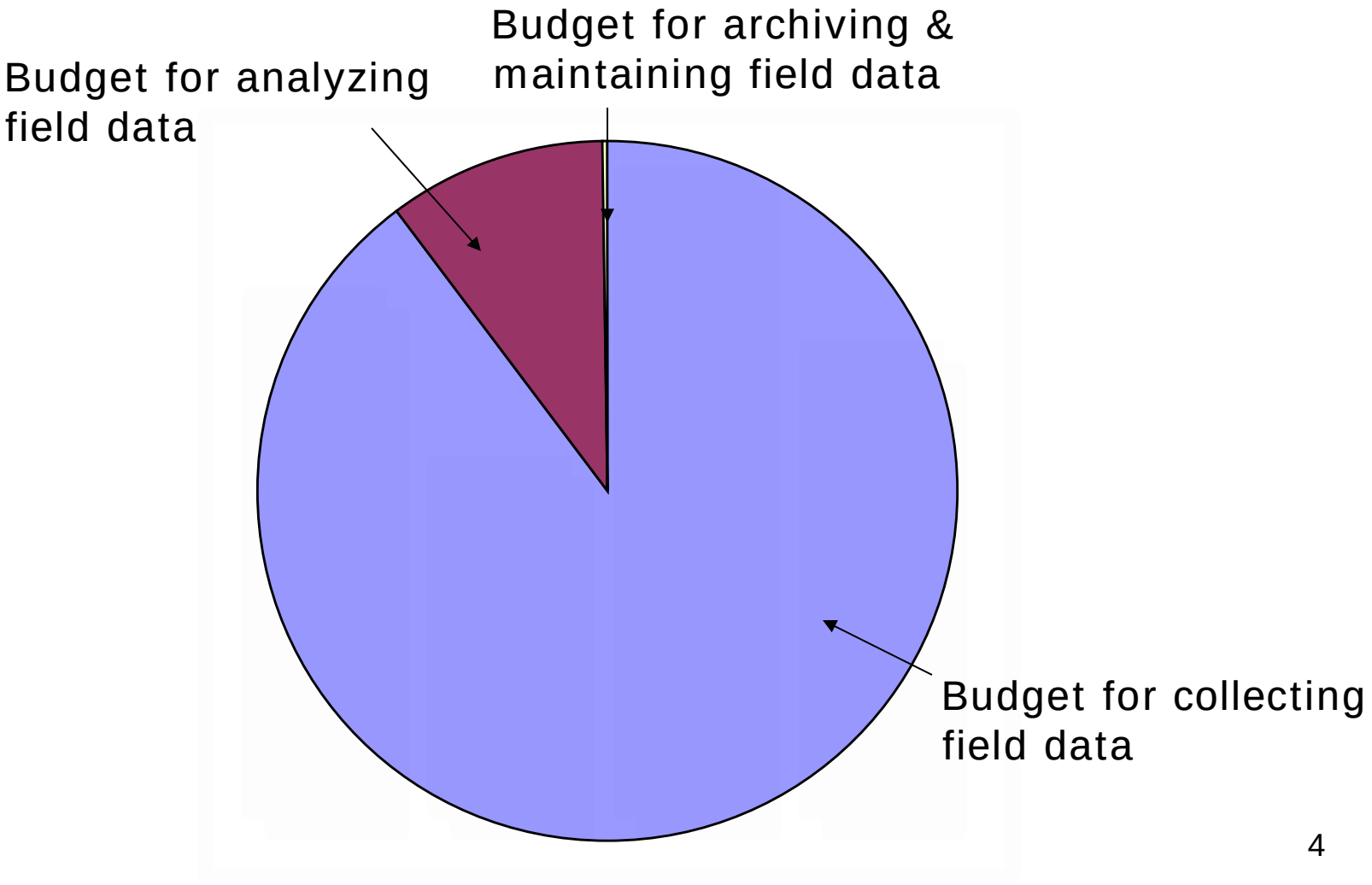
Outline

- Background
- Caveats
- Survey of existing ATD and meteorological data sets, and groups maintaining these data sets
- Necessary components of a good data set
- Common problems with data sets
- Recovery of data sets

Background

- Recommendation of the OFCM Joint Action Group on ATD
  - "High quality ATD data sets that exist from field research studies from the past 30+ years comprise the only existing basis for evaluation of ATD models"
    - At risk for loss of institutional memory, many scientists who were involved in experiments have retired or are retiring
    - At risk for loss of data, even for experiments that were done just a few years ago
    - Diverse forms and formats - difficult to access
    - Limited conditions
    - Limited analyses
    - Essential for studies of uncertainty
    - Encompass continental to urban scales
    - Cost-effective means to achieve re-analysis, not re-reinventing the wheel
- On May 10, 2005, OFCM organized a workshop on the *Recovery of Existing Atmospheric Transport and Diffusion Data.*
- Representatives from NOAA (DOC), DOD, DOE, EPA, and NRC gave presentations on the data sets that they currently have. Prof. Sue Grimmond of Indiana University briefed on international data sets.
- This presentation in part is based on the conclusions of this workshop.

The Reality


Budget for archiving & maintaining field data
Budget for analyzing field data
Budget for collecting field data

Sample Compilation of Results from Past Model Performance Studies

	Suggested Acceptance Criteria	Data Base or Study Name										
		Prairie Grass (GPM)	Prairie Grass (VLSTRACK)	Indianapolis (HPDM)	AERMOD study (5 models[1])	Dense gas MDA (10 models[2])	Kit Fox (3 HEGADAS versions)	DP26 (4 models[3])	OLAD (3 models[4])	Urban 2000 (20 HPAC config.)	Urban 2000 (20 HPAC config.)	ETEX (46 models)
No. of Trials	N/A	44	44	89 (hr[5])	100s (hr[5])	41	52	14	11	18	18	1 (90 hr)
Reference[8]	Chang and Hanna (2004)	Chang (1998)	Chang (1998)	Hanna and Chang (1993)	Hanna et al. (2000)	Hanna et al. (1993)	Hanna and Chang (2001)	Bowers et al. (2004). Chang et al. (2003).	Chang et al. (2003)	Chang et al. (2005)	Chang et al. (2005)	Girardi et al. (1998)
Output Considered	Unpaired in space	Arc-max conc.	Arc-max conc.	Arc-max conc.	Arc-max conc.	Arc-max conc.	Arc-max conc.	Arc-max dosage	Arc-max dosage	Arc-max conc.	Paired in space and time	Paired in space and time
Systematic Bias	< 30%	10% under	60% over	5%	2% - factor of 3	30% - 50%	5% to 50%	< 35%	Factor of 2 - 3 under	0 to factor of 4 over (median: 50% over)	25% to factor of 4 over (median: 50% over)	0 to factor of 4 (median: 5%)
Random Scatter[6]	Factor of 2 -3	Factor of 2	Factor of 4	Not reported	Factor of 2 - 7	Factor of 2	Factor of 2	Factor of 3 - 4	Factor of 4 - 7	Factor of 2 - 9 (median: factor of 3)	Factor of 9 - 25 (median: factor of 13)	Factor of 4 - 11 (median: factor of 5)
Fraction Within a Factor of 2	> 50%	80%	30%	65%	10% - 80%	50%	50% - 90%	40% - 60%	25% - 50%	5% - 60% (median: 40%)	20% - 30% (median: 25%)	10% - 40%[7] (median: 30%)

Notes

[1] Include five data sets (OPTEX, Duke Forest, Indianapolis, Kincaid, and Lovett), and three models (AERMOD, ADMS, and ISC3).

[2] Dense gas Modeler's Data Archives (MDA) include six data sets (Burro, Coyote, Desert Tortoise, Goldfish, Maplin Sands, and Thorney Island), and ten models (DEGADIS, HEGADAS, HGSYSTEM, SLAB, AIRTOX, CHARM, FOCUS, GASTAR, PHAST, and B&M Nomogram).

[3] HPAC, VLSTRACK, CALPUFF, and D2-PUFF.

[4] HPAC, VLSTRACK, and CALPUFF.

[5] Each hour is considered one trial.

[6] Random scatter is expressed in terms of a factor of N of the mean.

[7] Note that out of all the observation-prediction pairs that have been filtered by Girardi et al. (1998), 34% of the observed values are equal to zero. Therefore, a model predicting zero everywhere would have 34% of data that coincide with the observed values.

[8] References in Notes page.

Survey of ATD Data Sets - Caveats

- The data sets (~ 100) survey provided here is by no means exhaustive. There will be holes and any inputs are welcome! For example, the survey did not consider any weapon test programs, and Asian and Australian data sets.
- Some data sets may be maintained at multiple locations and *in multiple versions,* sometimes even *under different names.*
- *Complete (raw)* data for some data sets are only in hard copy or in antiquated media, or may have already been lost.
- Data availability can be unknown, in separate data files, or in web-based relational database management system (*e.g.*, Oracle).

Summary of Data Sets at HQ/ARL NOAA

- Air Resources Laboratory (ARL) Data Archive of Tracer Experiments and Meteorology (DATEM)
  - http://www.arl.noaa.gov/datem
  - ACURATE (Atlantic Coast Unique Regional Atmospheric Tracer Experiment)
  - ANATEX (Across North America Tracer Experiment)
  - CAPTEX (Cross Appalachian Tracer Experiment)
  - INEL74 (Idaho Falls National Engineering Laboratory, Kr^{85} in Midwestern U.S.)
  - OKC80 (regional-scale tracer experiment near OKC)
  - METREX (Metropolitan Tracer Experiment)

Information courtesy: Drs. Roland Draxler and Ray Hosker, NOAA

Summary of Data Sets at ATDD/ARL, NOAA

- Atmospheric Turbulence and Diffusion Division (ATDD)
  - Oak Ridge Reservation study, 1948-1952. Mainly wind fields in complex terrain, but some dispersion (balloons, smoke).
  - Radar-tracked tetrahedral balloons (tetroons) over LA (1963, 1969, 1973), Atlantic City (1965), NTS (1965-1966), NYC (1965), OKC (1971).
  - St. Louis. EPA-NOAA.
  - Denver "Brown Cloud". NOAA-ETL (old WPL).
  - Reactor dispersion studies. NRC.
  - Convective dispersion studies at Boulder Tower.
  - ASCOT (Atmospheric Studies in Complex Terrain). Complex terrain dispersion in northern CA, western CO, eastern CO, east TN. DOE.
  - Lab dispersion studies (wind tunnel and water channel), EPA-NOAA.
  - MVP (Model Validation Program). Cape Canaveral and Vandenberg AFB (1995-1997).
  - Joint Urban 2003. Oklahoma City.

Information courtesy: Drs. Roland Draxler and Ray Hosker, NOAA

Summary of Data Sets at FRD/ARL, NOAA

- Field Research Division
  1. Rush Valley Experiment (RVX), Tooele, UT (1987)
  2. Utah County Tracer Experiment, Provo, UT (1989)
  3. State of Idaho Plume Study (1993)
  4. Dipole Pride 26 (DP26), Nevada Test Site, NV (1996)
  5. Overland Alongwind Dispersion (OLAD), Dugway Proving Ground, UT (1997)
  6. San Luis Obispo (SLO) Coastal Study, San Luis Obispo, CA (1986)
  7. AMADEUS, Red Bluff, CA (1987)
  8. Kennecott Copper, SLC, UT (1990)
  9. Long-Range Overwater Diffusion (LROD), Kauai, HI (1993)
  10. Model Validation Program (MVP), Cape Canaveral, FL, and Vandenberg AFB, CA (1995-1997)
  11. Across North America Tracer Experiment (1987)
  12. Measurement of Haze and Visual Environments (MOHAVE), Laughlin, NV (1991)
  13. Big Bend Regional Aerosol and Visibility Study (BRAVO), Big Bend National Park, TX (1999)
  14. Princeton TOKAMAC, Princeton, NJ (1988)
  15. Full Scale Plume Study (FPSP), Reno, NV (1984)
  16. Cross Appalachia Tracer Experiment (CAPTEX), Southeastern US (1983)
  17. Small Hill Impaction Study (SHIS), Farmington, NM (1982)
  18. Idaho Field Experiment (IFX), INEL, ID (1981)
  19. Liquid Metal Fast Breeder Reactor (LMFBR), Clinch River, TN (1974)
  20. Big Southern Elevated Plume Study, INEL, ID (1974)
  21. NOAA's Arc I, ID (1972)
  22. NOAA's Arc II, ID (1977)

Information courtesy: Dr. Kirk Clawson, ARLFRD, NOAA

Summary of Data Sets at LLNL

- Dense Gas Data Sets
  - Burro (Naval Weapons Center, China Lake)
    - LNG, June-Sept 1980
  - Coyote (China Lake)
    - LNG, Rapid Phase Transition, summer and fall 1981
  - Desert Tortoise (Frenchman Flat, Nevada Test Site)
    - Ammonia, Aug, Sept 1983
  - Eagle (FF, NTS), Sept-Nov 1983
    - Nitrogen Tetroxide
  - Falcon (FF, NTS) June-Aug 1987
    - LNG
  - Goldfish (FF, NTS) Summer 1986
    - Anhydrous hydrofluoric acid
- Building 170 experiment
  - Tracer release around building

Information courtesy: Dr. Marty Leach, LLNL

Summary of Data Sets at GMU

http://camp.gmu.edu/data_resources_overview.html

1. Dense Gas MDA (Modeler's Data Archives, including Burro, Coyote, Desert Tortoise, Goldfish, Maplin Sands, and Thorney Island data sets)
2. Prairie Grass
3. Hanford Kr^{85}
4. Kit Fox
5. DTRA Phase I
6. DP26 (Dipole Pride 26)
7. OLAD (Overland Alongwind Dispersion)
8. MVP (Model Validation Program)
9. Ventura
10. Pismo Beach
11. Cameron
12. Carpinteria
13. LROD (Long-Range Overwater Diffusion)
14. MADONA (Meteorology And Diffusion Over Non-Uniform Areas)
15. ACURATE (Atlantic Coast Unique Regional Atmospheric Tracer Experiment)
16. ANATEX (Across North America Tracer Experiment)
17. METREX (Metropolitan Tracer Experiment)
18. CAPTEX (Cross Appalachian Tracer Experiment)
19. ETEX (European Tracer Experiment)
20. INEL74
21. OKC80
22. Birmingham
23. Joint Urban 2000 (Salt Lake City)
24. Joint Urban 2003 (Oklahoma City)
25. MUST (Mock Urban Setting Test)
26. EMU
27. BARREL
28. LA 2001
29. Barrio Logan
30. Macdonald
31. SMEDIS
32. TRAPOS (Optimization of Modeling Methods for Traffic Pollution in Streets)
33. REDIPHEM (Review and Dissemination of Physical Effects Models)
34. FLADIS
35. Chesapeake Bay 2001
36. Kincaid
37. Bull Run
38. Indianapolis
39. Clifty Creek
40. Tracy
41. Martins Creek
42. Westvaco
43. SARMAP (San Joaquin Valley Air Quality Study, Regional Meteorological and Air Pollution)
44. LMOS (Lake Michigan Ozone Study)
45. OTAG (Ozone Transport Assessment Project)

Have actual data for most data sets.

Summary of Data Sets at GMU

http://camp.gmu.edu/data_resources_overview.html

		Prairie Grass	DP26	OLAD	LROD	MADONA	Burro	Coyote	Desert Tortoise	Goldfish	Maplin Sands	Thorney Island	Hanford Kr85	Kit Fox	DSWA Phase I	MUST	Ventura	Pismo Beach	Cameron	Carpinteria	MVP	CAPTEX	ETEX	ACURATE	METREX	Kincaid	Bull Run	Indianapolis	Clifty Creek	Tracy	Martins Creek	Westvaco	Urban 2000 (SLC)	Urban 2003 (OKC)	Los Angeles	Barrio Logan
Parameters	Sub-Categories																																		
Source Geometry	Point	x	x			x			x	x			x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Line			x	x																												x			
	Area						x	x			x			x																						
	Volume											x			x																					
Source Term	Instantaneous		x	x	x	x						x			x	x					x												x			
	Finite duration with constant (or slowly varying) rate	x				x	x	x	x	x	x		x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Finite duration with varying rate																																			
Source Elevation	Ground (0 to ~50 m)	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x								x	x	x	x
	Elevated (~50 m to atmospheric boundary layer)			x	x															x	x					x	x	x	x	x	x	x				
	High-altitude (beyond atmospheric boundary layer)																																			
Source Dynamical Effects	Neutrally buoyant	x	x	x	x	x							x		x	x	x	x	x	x	x	x	x	x	x								x	x	x	x
	Buoyant																									x	x	x	x	x	x	x				
	Dense						x	x	x	x	x	x		x																							
	Momentum (jets)								x	x																											
Source Location	Stationary	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Moving			x	x																															
Terrain	Flat (grass, crops, desert, forest, etc.)	x	x	x					x	x		x	x	x	x	x						x	x	x	x	x										
	Complex (mountains, valleys, etc.)		x	x		x																x	x	x			x		x	x	x	x				
	Coastal (land-water interface)																x	x	x	x	x															
	Overwater				x		x	x			x																									
	Urban															x									x			x					x	x	x	x
Time of Day	Day (neutral to unstable)	x	x	x	x	x	x	x	x	x	x		x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x
	Night (neutral to stable)	x	x			x							x	x	x	x			x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Transition			x		x						x									x	x	x	x	x	x	x	x	x	x	x					
Prevailing Weather	Steady (persistent)	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Frontal passage					x															x	x	x	x	x	x	x	x	x	x	x					
Meteorological Measurements	Surface (including towers)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Radiosondes (including tethersondes)		x	x	x	x										x				x	x	x	x	x	x	x	x	x	x	x	x	x	x			
	Remote sensing (profilers, sodars, etc.)		x	x		x										x					x					x	x	x				x	x			x
	Sonic anemometers (turbulence data)		x	x		x									x	x																x	x			x
	Aircraft				x												x	x	x		x															
Sampler Distribution	Rows (includng rectangular arrays and lidar scans)	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x						x	x	x		x			x	x	x	x
	Randomly spaced																					x	x	x	x			x		x	x				
	Aircraft or moving van				x																x															
Sampler Downwind Distance	Near source (~ < 2 km)	x					x	x	x	x	x	x	x	x	x	x				x															x	x
	Meso-gamma scale (~ 2 – 20 km)		x	x		x											x	x	x						x	x	x	x	x	x	x	x	x		
	Meso-beta scale (~ 20 – 200 km)				x																x															
	Meso-alpha scale (~ 200 – 2000 km)																					x	x	x												
Concentration Measurements	Time averages	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	High frequency time series (e.g., ~ 1 Hz)		x	x	x	x	x	x	x	x		x		x	x	x					x												x			
	Repetition of ensemble members					x									x	x																				
	Deposition																																			
Scale	Full scale field	x	x	x	x	x	x	x	x	x	x	x	x		x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
	Reduced scale field													x		x																					
	Laboratory																																			

Summary of Data Sets at EPA

- Field experiments
  - Buoyant plume, no downwash
    - Flat terrain: Prairie Grass, Kincaid, Baldwin, Indianapolis (urban), St. Louis (urban)
    - Complex terrain: Clifty creek, Tracy, Martins creek, Lovett, Westvaco
  - Building wake
    - Bowline point, Alaska North Slope, Millstone Nuclear Power Plant, Duane Arnold Energy Center, Organic Cooling Reactor
- Fluid Modeling Facility (mostly neutral boundary layer)
  - Dense gas small area sources, jets, and line sources
  - Isolated idealized terrain
  - Two-dimensional valley study
  - Urban building arrays
  - Idealized twin high-rise buildings
  - World Trade Center
  - Concentration fluctuations

Information courtesy: Dr. S.T. Rao, NOAA/EPA

Summary of Data Sets at DPG

- Complete data sets
  - LROD (1993)
  - DSWA Phase I (1996)
  - Dipole Pride 26 (1996)
  - OLAD (1997)
  - MUST (2001)
  - Joint Urban 2003 (2003)
    - https://ju2003-dpg.dpg.army.mil
    - Single-point data warehouse (DBMS) for all JU2003 data
    - Web interface to download subsets of data
    - Redundant systems
    - Quarterly audits
    - Will be maintained by DPG for several years, subject to funding
- Partial data sets
  - Program Wind (1987)
  - Concentration Fluctuation (1991-1995)
  - MADONA (1992)
  - Kit Fox (1995)
  - GRADE (1997)
  - Urban 2000/VTMX (2000)
  - Pentagon Shield (2004)

Information courtesy: Mr. Donny Storwold, WDTC, DPG

Summary of Data Sets at NRC

- Building Wake Effects at Millstone Nuclear Power Plant (CT)
- Low Wind Speed Data at River Bend (LA)
- Shoreline Environment Atmospheric Dispersion Experiment at Kewaunee (WI) – SEADEX
- Idaho Field Experiment (ID)

Information courtesy: Ms. Jocelyn Mitchell, NRC

Summary of CBD Data Sets at NSWC, Dahlgren

- Model Validation Database (web-based DBMS)
  - Contains data describing larger releases of an agent - more typical of an intentional release
  - MUST validation data being added
  - Urban 2000 data (Salt Lake City)
  - Joint Urban 2003 data collected (Oklahoma City)
  - Fort Wayne data sets collected
- Surface Evaporation Database (web-based DBMS)
  - Contains data describing the three scales of surface evaporation trials
    - Field trials
    - Wind tunnel
    - Bench scale
  - Data describing the behavior of droplets on various surfaces
  - Supported by the Agent Fate literature survey and analysis task

Information courtesy: Mr. Gaurang Dävé, NSWC, Dahlgren Division

Summary of European Data Sets, Urban

Canyon	Neighborhood	City	Region	Name	Location
X				TRAPOS	Nantes
X				TRAPOS	Berlin
X				TRAPOS	Hannover
X				TRAPOS	Copenhagen
X				GOTE-DUCC	Goteborg
	X	X			Copenhagen*
	X	X	X	ESCOMPTE	Marseille
X	X	X		BUBBLE	Basel
		X	X	PACIFIC 2001	Vancouver
		X	X	PACIFIC 1993	Vancouver
					Lillestrom
	X	X			Birmingham

Information courtesy: Dr. Sue Grimmond, Indiana University

Summary of European Data Sets,
Coastal (top), Complex Terrain (middle), and Dense Gas (bottom)

Name	Location
PACIFIC 1993, 2001	Vancouver, Lower Mainland BC, Canada
BP International	England (Industrial)
Teruel	Spain (Industrial)
ESCOMPTE	Marseille, France

Name	Location
MAP: Riveria	Switzerland
LTEX	New Zealand
ETEX	France, West Europe (long range transport)
ESCOMPTE	Marseille, France
BP International	England

HF	C_3H_8	NH_3	Name	Location
X			URAHFREP	Porton Down, England
	X		REDIPHEM	Lathen
		X	FLADIS (FLADEXP)	Sweden

Information courtesy: Dr. Sue Grimmond, Indiana University

Wind Tunnel Studies at University of Hamburg

http://www.mi.uni-hamburg.de/technische_meteorologie/windtunnel/

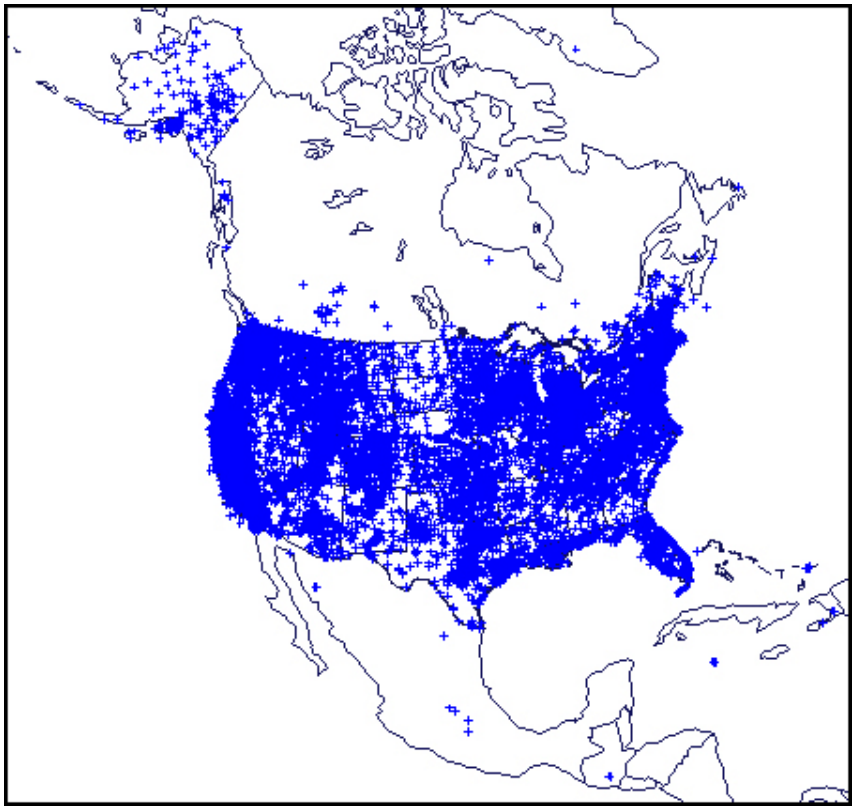
- Led by Prof. Dr. Michael Schatzmann and Dr. Bernd Leitl
  - Urban dispersion
  - Heavy gas dispersion
  - Buoyant plumes
- Many of these wind tunnel studies compliment full-scale field studies in U.S. and Europe (*e.g.*, Kit Fox)

Information courtesy: Dr. Sue Grimmond, Indiana University

Summary of Meteorological Data Sets

- Different from meteorological and ATD data sets
  - In addition to special meteorological data collected during ATD field experiments, there are many routine weather observations. Many global and mesoscale meteorological models are also run operationally. These *routine observations* and *model products* are systematically archived.
- National Climatic Data Center (NCDC)
  http://www.ncdc.noaa.gov
- Forecast System Laboratory (FSL)
  http://www.fsl.noaa.gov/data/
  http://www-sdd.fsl.noaa.gov/MADIS/
- National Center for Atmospheric Research (NCAR)
  https://cdp.ucar.edu/
- European Centre for Medium Range Forecast (ECMWF)
  http://www.ecmwf.int/products/data/

MADIS (Meteorological Assimilation Data Ingest System) Surface Mesonet System



- Current mesonet sites = 16,667
- Current mesonets = 98
- Largest mesonets
  - AWS Convergence Tech (= 6825)
  - Citizen Weather Observing Program (= 3251)
  - National Interagency Fire Center's (NIFC) Remote Automated Weather System (RAWS)

Information courtesy: Dr. Sharon LeDuc, NDCDy

Necessary Components of a Good Data Set

- Metadata
  - Site descriptions (terrain, vegetation, soil, imagery, maps, *etc.*)
  - Instrumentation (accuracy, threshold, QA/QC procedures, *etc.*)
- Raw data
  - Meteorological data
    - ~20 Hz to twice-daily
    - Surface and upper air
    - In situ and remote
  - Source data
  - Concentration data (~10 Hz to 3-hr)
  - Terrain elevation data
  - Building data
- Processed data
  - Reduced (*e.g.*, averaged) data
  - Summary tables
  - Modeler's Data Archives, minimum set to run models

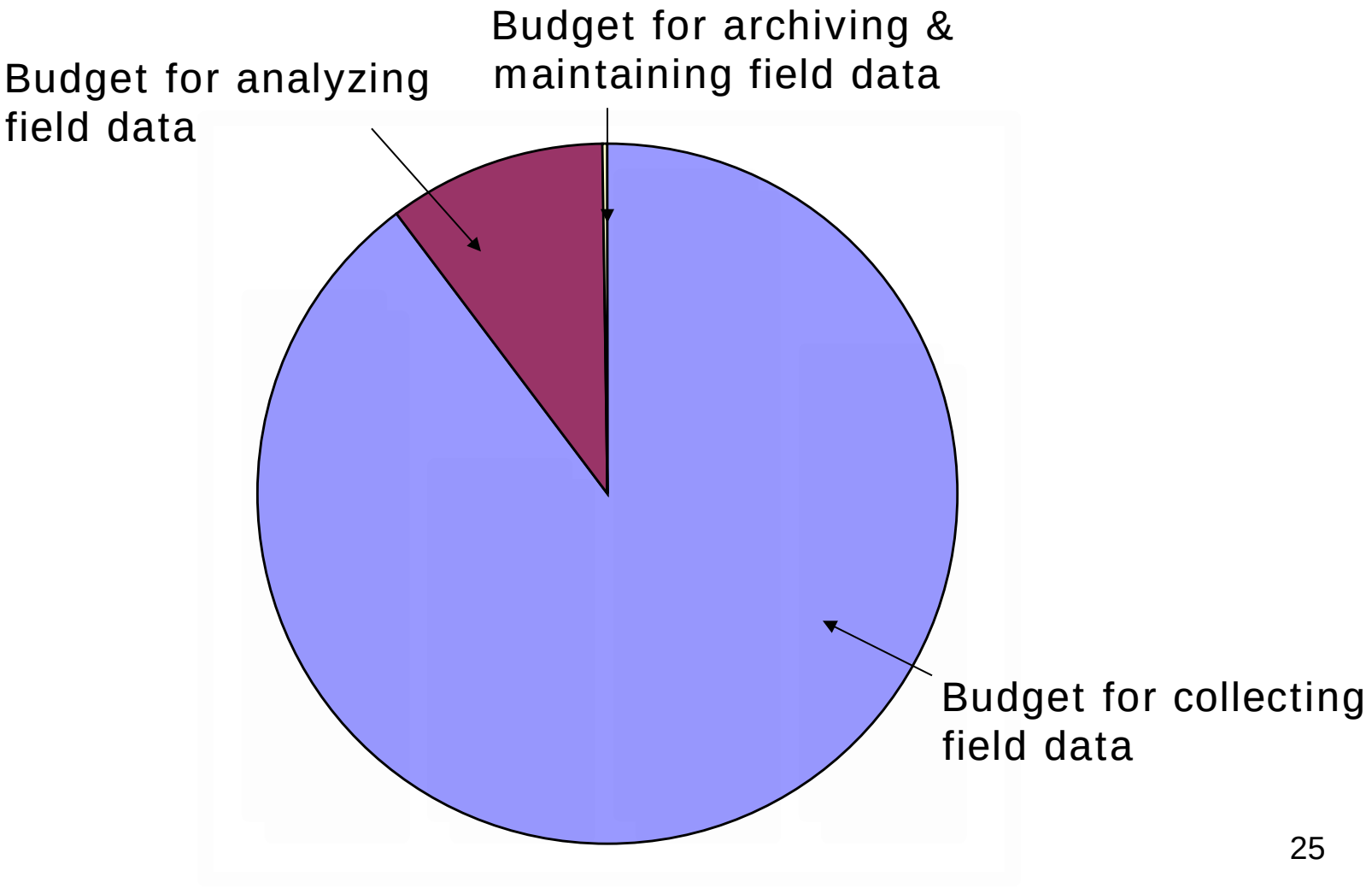
Common Problems with Data Sets

- Lack of adequate metadata (site descriptions, instrumentation accuracy and threshold, *etc.*)
- Lack of adequate meteorological data (characteristic of many early weapon test programs)
- Inconsistent, confusing units, time zones, coordinates, and missing data indicator, usually a result of inadequate experiment planning
  - m/s, knots, and mph
  - LST, LDT, and UTC
  - Lat/lon, and UTM (which datum, WGS84 or Clarke 1866?)
  - -9999, -999, and N/A
- Lack of or unknown data QA/QC
- Diverse media and formats
  - In some cases, reports are the only available records of the experiment and no electronic data are available. The data for very few of such data sets (*e.g.*, Prairie Grass) have been digitized.
  - In some cases, data are stored in antiquated media such as strip charts, 9-track tapes, and 5.25” floppy disks.
  - In some cases, data are in uncommon binary formats (*e.g.*, 40-bit word)
  - ASCII is probably the best choice because it can be easily read by all types of applications. (Limit on disk storage is basically not an issue.)

Recovery of Data Sets, A Suggestion

- Start a pilot project that
  - Identifies good-quality data sets *useful* to ATD modeling.
  - Creates a master list of these data sets with good descriptions.
  - Prioritizes data sets based on, for example
    - Can it be done quickly?
    - Are SMEs available?
    - How much does it cost?
    - What are the benefits?
    - Are the data sets unique or rare, thus filling a data gap?
  - Collects data sets (with all the necessary components mentioned above) from various sources.
  - Performs additional QA/QC and post-processing, if necessary.
  - Creates a data server hosting data sets.
    - A data server can be
      - A portal to other existing sites
      - An ftp site with individual data files and reports
      - A web-based server with full-blown DBMS (*e.g.*, JU2003)
    - Need to identify subject matter experts (SMEs) for each data set
    - These SMEs are notified whenever data are requested
- Need stable funding source!!!!!

The Reality


Budget for archiving & maintaining field data
Budget for analyzing field data
Budget for collecting field data

A Better Scenario