REVIEW OF AIR DATA SYSTEMS

prepared for

The U.S. Environmental Protection Agency

Contract No. 68-01-3308

Task Order 68-01-3094

prepared by

Index Systems, Inc.

One Broadway

Cambridge, Massachusetts 02142

December, 1976

REVIEW OF AIR DATA SYSTEMS

Prepared For

THE U.S. ENVIRONMENTAL PROTECTION AGENCY

Contract No. 68-01-3308
Task Order 68-01-3094

Prepared by

INDEX SYSTEMS, INC.
One Broadway
Cambridge, Mass.

December 1976

TABLE OF CONTENTS

INDEX OF EXHIBITS

I. MANAGEMENT SUMMARY

I. MANAGEMENT SUMMARY

During the period from January 19, 1976 through July 30, 1976, Index Systems conducted a review of the major automated data processing (ADP) requirements and systems operated on behalf of the air pollution control programs of the U.S. Environmental Protection Agency (EPA). This report presents (1) the study team's findings and conclusions regarding air program information needs, (2) the efficiency and adequacy of current ADP usage to support these needs, and (3) recommended long-term policies and short-term action steps to reduce ADP costs and improve the quality of ADP services in support of air program data requirements.

This management summary presents the key results contained in the report. It is divided into five sections, as follows:

- Project Objectives and Scope
- Project Approach
- Summary of Findings
- Summary of Recommendations
- Report Organization

1. PROJECT OBJECTIVES AND SCOPE

The main objective of this study has been to examine the legislative and programmatic requirements of the air program and to evaluate the degree to which current EPA air computer systems satisfy these requirements. Particular emphasis was to be placed upon evaluating the use of separate national and state data systems, upon evaluating the need for separate program and research air monitoring data systems, and upon analyzing the relative merits of current system design philosophies as opposed to more centralized and more distributed design approaches.

The scope of the study included four groups of systems, as follows:

Systems operated by the Office of Air Quality, Planning and Standards (OAQPS) including:

- NEDS - National Emissions Data System
- SAROAD - Storage and Retrieval of Aerometric Data
- SOTDAT - Source Test Data System
- HATREMS - Hazardous and Trace Element Inventory System
- SIPS - State Implementation Plans System
- QAMIS - Quality Assurance Management Information System
- EDS - Energy Data System
- FPC (67) - Federal Power Commission Data
- SIEFA - Source Inventory and Emission Factor Analysis Program
- WSAP - Weighted Sensitivity Analysis Program
- REPS - Regional Emission Projection System
- CAASE - Computer-Assisted Area Source Emission Gridding Procedure
- PRMS - Plans Review Management System

- Data systems operated by the Office of Research and Development (ORD) to support air program research activities including:

  RAPS - Regional Air Pollution Study

  CHESS - Community Health and Environmental Surveillance System

  CHAMP - Continuous Health Air Monitoring Program

  NASN - The filter bank system of non-criteria pollutant data

- The Compliance Data System (CDS) operated by the Office of Enforcement (OE)
- The Comprehensive Data Handling Systems (CDHS) operated by state and local agencies

Each system or group of systems was reviewed with regard to its purpose, usage, efficiency, functional adequacy and impact upon air program activities.

2. PROJECT APPROACH

The activities undertaken to achieve the project objectives consisted of five tasks, as follows:

(1) Analyze User Needs

During this task the Index project team reviewed and documented the overall national air pollution control strategy and characterized the data needs associated with each major program activity.

These information requirements were used during the subsequent tasks to evaluate the adequacy of current system support and to assess proposed changes in systems scope, design and operations.

(2) Evaluate System Functional Capabilities

During this task the project team reviewed the functional capabilities of the EPA air data systems and evaluated the adequacy of these capabilities in satisfying identified user information needs.

(3) Evaluate System Technical Characteristics and Costs

During this task the project team examined the technical efficiencies of the various air data systems. This examination covered four areas, as follows:

- Efficiency of systems implementation and operation on the EPA computer facilities
- Conformity of systems design and documentation with sound ADP principles
- Data volumes and data base management procedures employed
- Current and future system costs for each system including personnel, contractor and computer expenditures

As a result of this task the project team linked the identified system functional requirements with technical system design and operation efficiency and costs. Based on this information the project team prepared preliminary system recommendations regarding system modifications or enhancements to improve the cost effectiveness of air system capabilities in meeting high priority user needs.

(4) Evaluate System Management Procedures

During this task the project team analyzed the adequacy of current ADP system management practices and procedures. Particular emphasis was placed upon the cost effectiveness of current system support organization and responsibilities, the adequacy of current system personnel resources and the responsiveness of system managers to the user community.

(5) Develop Recommended Changes

Based on the preceding tasks, the project team developed recommendations to improve air data system capabilities, operations and management. These changes and the continued operation of current systems were cost-justified on the basis of legislative and programmatic requirements as determined during Task 1. In addition, expected costs and resource requirements for continued operations, enhancements, and implementation of recommendations as well as expected cost savings from design efficiencies and system consolidation were projected.

3. SUMMARY OF FINDINGS

The study team reviewed the adequacy of current and anticipated future ADP systems in satisfying air program information needs. In addition, the study team assessed the technical efficiency of air data systems in providing required services.

(1) EPA's Overall Air Pollution Abatement Strategy Involves a Shift in Day-to-Day Operating Responsibilities to State and Local Agencies.

Prior to the creation of EPA, most federal air pollution control programs were conducted in North Carolina by the organization which later became EPA's air program office. With the formation of EPA and with the passage of the Clean Air Act, the Office of Air Quality Planning and

Standards (OAQPS) became responsible for developing a national air pollution abatement program and for assisting state and local agencies in developing and implementing regulations and control strategies to meet nationwide standards. In recent years, as more regional and state personnel have received training in air pollution control methods, the relative responsibilities of Headquarters and field personnel have been changing. Exhibit 1-1, on the following page, summarizes the current program functions. Headquarters is responsible primarily for overseeing national progress, for developing new strategies for the control of criteria and non-criteria pollutants, and for developing new tools for use in reducing current pollution levels. EPA regions are responsible primarily for overseeing and monitoring state progress in achieving air pollution abatement objectives. State and local agencies are responsible primarily for developing and enforcing regulations that are part of the State Implementation Plans (SIP). New programmatic functions, such as National Emissions Standard for Hazardous Air Pollutants (NESHAPS) regulations or non-significant deterioration regulations, are developed by EPA but may be delegated to state agencies.

Exhibit 1-1
U.S. Environmental Protection Agency
Summary of Federal and State Air Program Responsibilities

FUNCTION	EPA	STATE AGENCIES
Standards Development	Establish national ambient standards Establish NESHAPS Revise standards Analyze new pollutants Establish mobile source emission standards Develop new source performance standards	Establish state standards Revise standards
Plan Development and Implementation	Review SIPS and SIP progress Test mobile source emissions Conduct fuels and vapor recovery programs Review enforcement activity Provide technical assistance Develop and delegate NSD and AQMA programs Conduct or delegate NESHAPS program	Develop, evaluate and revise SIPS Develop TCP's Conduct new source reviews Write permits Conduct inspections and undertake enforcement actions Monitor ambient air quality
Special Activities	Perform health effects research Develop new control strategies Develop simulation models ESECA	Land use planning

(2) State and Local Agencies Require Access to Detailed, Local Ambient and Emissions Information in Order to Conduct Abatement Control Functions.

In order to carry out their day-to-day operational functions, state and local agencies require comprehensive ambient air quality and emissions information to assure compliance with standards and to support source-related control activities. Most emissions and air quality data used by state and local agencies are provided by local monitoring stations and through industrial or governmental source inspections. Increasingly, however, states have begun to use transportation and demographic data to deal with mobile source-related pollutants, such as CO and oxidants, in geographic areas with significant mobile source problems. Nonetheless, the majority of local data gathered and used, at present, is TSP and SO_2 monitoring information.

(3) EPA Requires National Trend Data to Support Program Management Functions and Specialized Data to Support Program Planning.

EPA planners and program management personnel require data on national air quality trends in order to assess nationwide performance in meeting Clean Air Act objectives. In addition, EPA research and program personnel require very accurate and specialized ambient and industrial data to assess health effects of criteria and non-criteria pollutants, and to develop new pollution control tools to expedite abatement activities. The Standing Air Monitoring Work Group (SAMWG) task force has determined that a much smaller number of trend stations will be needed in the future to support EPA Headquarters program management functions, that a slightly larger data base of air monitoring data will be needed by regional offices, and that local data will be needed by state and local agencies.

(4) <u>Current Air Data Systems and ADP Expenditures Concentrate Primarily Upon Maintaining Detailed National Data Bases of Ambient Information</u>.

Approximately one-half of air program expenditures for data processing support and for related contractor and personnel activities relate to the assembly, storage and review of raw ambient data produced by state and EPA monitoring networks. Exhibit 1-2, on the following page, summarizes the current ADP expenditures in support of air program activities for fiscal years 1975 and 1976. The SAROAD, RAPS, CHESS/CHAMPS and NASN expenditures of $1.55 million in 1976 are all for storing and manipulating

Exhibit 1-2
U.S. Environmental Protection Agency
Air Data System Costs

	1976	1975
SAROAD (includes regions)	$ 903,000	$ 606,242
NEDS	150,000	260,131
CDS	105,000	48,000
RAPS	250,000	117,600
CHESS/CHAMPS	347,000	289,756
EDS	183,000	18,000
CDHS	300,000	Not Available
NASN	54,000	89,000
Other	834,000	875,378
TOTAL	$ 3,126,000	$ 2,215,107

ambient air quality data. These costs are in addition to state-incurred costs for CDHS and other state-operated systems to manage local ambient data. This emphasis upon ambient air data acquisition and manipulation reflects a policy adopted in prior years by OAQPS whereby EPA maintained a central data bank of national ambient air quality information. Recently, however, this policy has undergone revision and increasingly states have been encouraged to take responsibility for acquiring their own data banks of ambient data, as needed, to carry out abatement and enforcement functions.

(5) <u>The Instability of the UNIVAC Computer Has Created Problems for All Users of Air Data Systems</u>.

Since its installation in 1974, UNIVAC hardware problems at the National Computer Center (NCC) have created an unstable system environment which has inhibited timely data processing operations and has limited the accessibility of the information stored in the principle air data systems. Frequent system crashes and prolonged downtime have increased processing backlogs and have delayed software conversion and enhancement efforts. In recent months, actions have been initiated to correct hardware malfunctions. Although some improvements have been implemented, the mean time between failure is still below the NCC goal of 15 hours.

(6) <u>Currently Planned Changes in Air Data Systems Will Make Them Less Sensitive to UNIVAC Hardware Problems</u>.

EPA has undertaken substantial changes in the design of its major air data systems, NEDS and SAROAD, which had been converted inappropriately when transferred to the UNIVAC. These changes have been initiated in order to reduce the running time required to perform system updates and to facilitate report retrieval. Most changes are scheduled for completion by the end of fiscal year 1976. Current estimates suggest that these changes will reduce SAROAD operating costs by 50%. More importantly, however, these processing cost savings will greatly reduce the risk exposure of NEDS and SAROAD to hardware and operating system failures on the UNIVAC computer. For example, a typical SAROAD update, which currently takes 30 hours to execute, is estimated to take only six hours after the new changes are completed. This shortening of the update time requirements will reduce substantially the chances of the computer "crashing" while a SAROAD update is in progress and thereby will reduce proportionately the number of times SAROAD must be reupdated to correct for the computer "crash". As a result, SAROAD will be updated on a more regular basis than at present and users will be able to access the data base without the week- or month-long delays which have been experienced recently.

(7) <u>Data Handling Problems and Delays in State-submitted Data Contribute to Data Currency Problems</u>.

NEDS and SAROAD data handling and editing procedures require many separate steps and thereby delay the entry of current corrected data. In addition, poor interfaces and non-compatible edits between state and federal data systems complicate the data correction process. As a result, Headquarters, regional office and state personnel must interact extensively to track data submittals and corrections. However, various regional offices and state and local agencies have placed low emphasis upon submitting and correcting submitted data. Consequently, EPA air data system users have been hampered in analyzing some state-submitted air quality and emissions data because of data currency problems.

(8) <u>Proposed Changes in the Air Monitoring Strategy Will Reduce the Volume of State-collected Data Stored Centrally by EPA.</u>

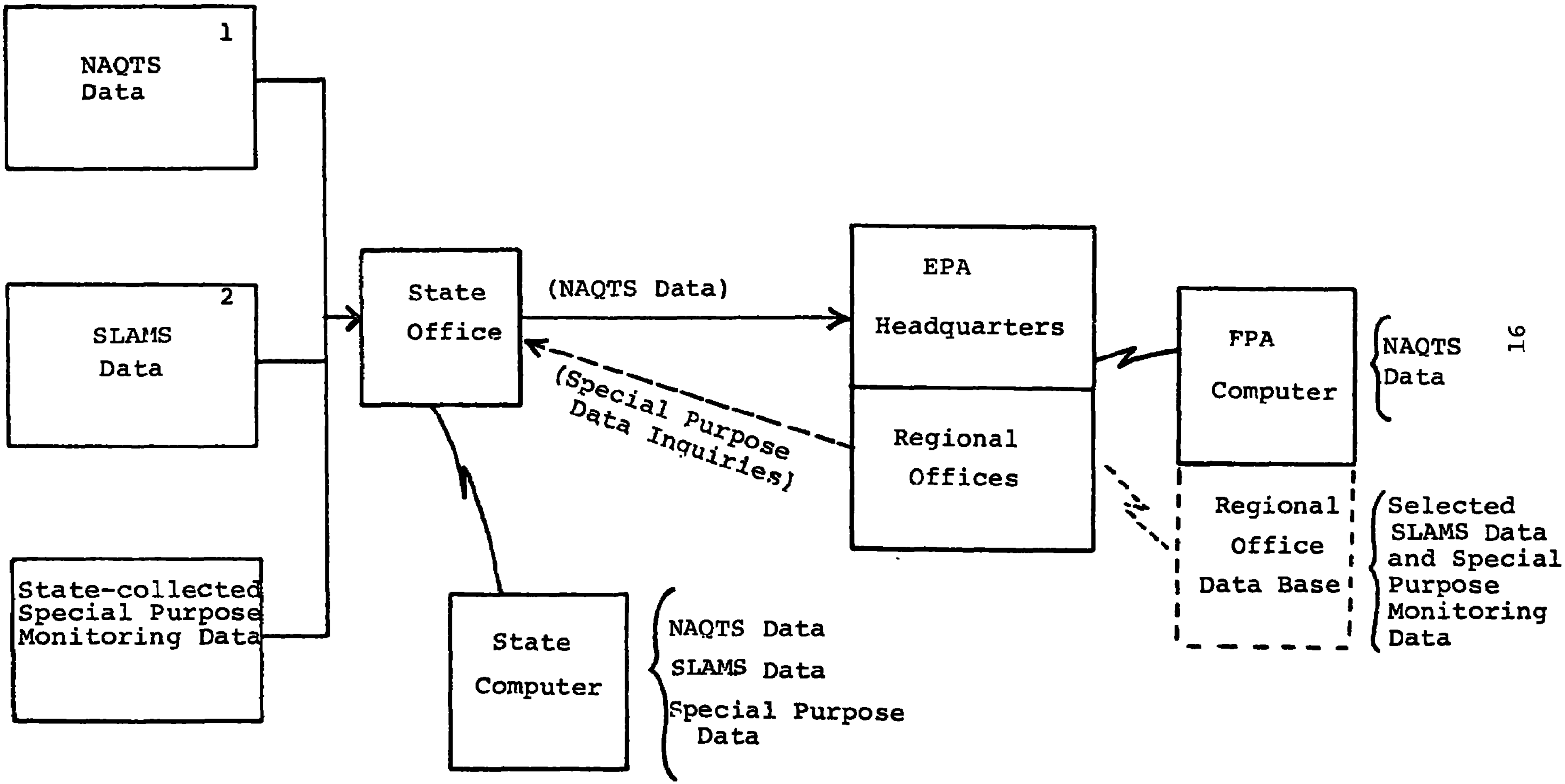
The Standing Air Monitoring Work Group (SAMWG) has undertaken a project to revise the requirements for the submittal of state-collected air monitoring data. Preliminary results from SAMWG indicate that EPA will reduce substantially the volume of data which it will require on a regular basis from states. State-collected air quality data will be grouped into three groups, as follows:

- National trend station data
- State trend station data
- Special purpose monitoring data

Exhibit 1-3, on the following page, illustrates the proposed air monitoring data reporting requirements associated with these three classes of data.

Under this proposed strategy a subset of the state-collected data, National Air Quality Trend Station (NAQTS) data, would be communicated to EPA regularly and stored on a central data base for analysis. SAMWG estimates that the number of monitoring sites reporting data regularly to EPA will be reduced by 80 percent. Some regional offices may elect to require some non-National Trend Station data reporting from states as may be defined in the future by SAMWG. In most cases, Regional Offices and others requiring summary statistics or raw data from non-NAQTS would obtain this information, as needed, from the state office which collects this special purpose data.

Exhibit 1-3
U.S. Environmental Protection Agency
Proposed Air Monitoring Data Reporting
Requirements



1 NAQTS - National Air Quality Trend Stations
2 SLAMS - State and Local Air Monitoring Stations

(9) <u>The Absence of Formal ADP Responsibilities and Procedures Has Lessened the Efficiency and Effectiveness of ADP Utilization in Support of the Air Program.</u>

Current EPA policies place operating responsibilities for national air data systems within the National Air Data Branch (NADB) in OAQPS. Although NADB contains the largest number of trained ADP personnel with experience in project management and development, current responsibilities do not reflect centralized control of system development projects within OAQPS. For example, the Energy Data System (EDS), which was not developed by NADB, will cost 10 times as much in ADP costs next year as the original feasibility study justified.

More broadly, ADP users of the National Computer Center's (NCC) UNIVAC computer are unaccustomed to managing or controlling their ADP utilization to conform to budget ceilings. Most system managers at the Research Triangle Park laboratories, for example, were unable to provide estimates of current- or future-year processing needs. Instead, research managers tended to view ADP budgets as interchangeable with program budgets and were not concerned, therefore, with ADP budgets since they intended to reprogram from research into ADP sub-allowances whenever necessary. Air program personnel indicated a greater awareness of ADP budgeting but did not have good accounting tools available whereby they could monitor and control actual UNIVAC computing expenditures. Regional ADP managers, by contrast, indicated the greatest sensitivity to NCC computing costs and demonstrated the most aggressiveness in developing methods to manage and control ADP expenditure levels.

4. SUMMARY OF RECOMMENDATIONS

EPA should initiate actions to shift towards a hybrid data processing philosophy for ambient air data and should complete currently-contracted system enhancements to reduce overall air data system expenditures.

(1) EPA Should Adopt a Hybrid Computer System Design Approach for Storing Air Quality Data in Conformity With the Revised Monitoring Strategy.

Costs will be reduced slightly and service will be improved if EPA adopts a hybrid system design philosophy for the management of state-collected air quality data. The hybrid system approach reflects the programmatic shift in monitoring strategy as being developed by the Agency's Standing Air Monitoring Work Group (SAMWG). This programmatic shift will require far less raw ambient data for Headquarters and regional use and will free states and local agencies to allocate more monitoring resources for special-purpose local projects. Exhibit 1-4, on the following page, contrasts the impact of a hybrid design approach with a utility and the current design approach. The hybrid and utility system alternatives would improve data handling and correction capabilities and would offer Headquarters and states flexible reporting and good data availability. By comparison, under the current system approach data handling and correction problems would continue to contribute to completeness and currency problems of EPA-maintained data as a result of the large volumes of data transferred and the relative distance and processing steps between the sources of data and the EPA users.

A hybrid approach would reduce slightly aggregate computer operating costs for EPA and state offices, yet would entail development and installation costs comparable to those needed to support the current design philosophy.

In addition, a hybrid system design would offer EPA lower technical and cost risks than the utility or current

Exhibit 1-4
U.S. Environmental Protection Agency
Comparison Summary of System
Design Alternatives

	Utility	Hybrid	Current
Functional Service	Improved data handling with good availability of state-collected air quality data	Improved data handling with good availability of monitoring trend data	Good data availability and reporting flexibility but poor data handling
Annual Operating Cost Impact	$625,000	$600,000	$800,000
Development and Installation	$600,000 - $1,000,000	$380,000	$330,000
Risk	High technical and cost risk	Low technical and cost risk	Moderate technical risk
Management Considerations	System management problems with state dependence upon EPA for ADP support	Less complex system management	Less dependence of EPA and states upon each other

approaches. Under a hybrid design EPA would maintain a much smaller data base consistent with the SAMWG data reporting requirements. This smaller data base would be less sensitive to hardware problems. Similarly, the cost risk would be reduced since the majority of data processing would be controlled by the states who, in turn, would be free to decide and pay for whatever data they needed for local purposes. A hybrid design would provide EPA with a system which is easy to manage and which would not entail heavy dependence of states on EPA-provided ADP support.

(2) <u>EPA Should Initiate Five Major Action Steps During the Next 2-3 Years in Order to Establish a Hybrid System for the Management of Air Quality Data</u>.

EPA should phase the implementation of the hybrid system approach over a 2-3 year period after EPA acceptance of the SAMWG data reporting recommendations. The phasing of this approach is necessary to facilitate management control and to insure a smooth and non-disruptive phasing of individual states into the hybrid data reporting strategy. The implementation process should be divided into five steps, as follows:

- EPA should define the specific data reporting requirements and procedures needed to support the SAMWG-recommended changes in air monitoring strategy. These definitions should include procedures for regional offices to acquire and maintain non-NAQTS data, as needed. Furthermore, EPA should develop procedures for support of those states without air quality data systems.

- OAQPS and the regions should select three states as prototypes to test the new monitoring strategy data reporting requirements.

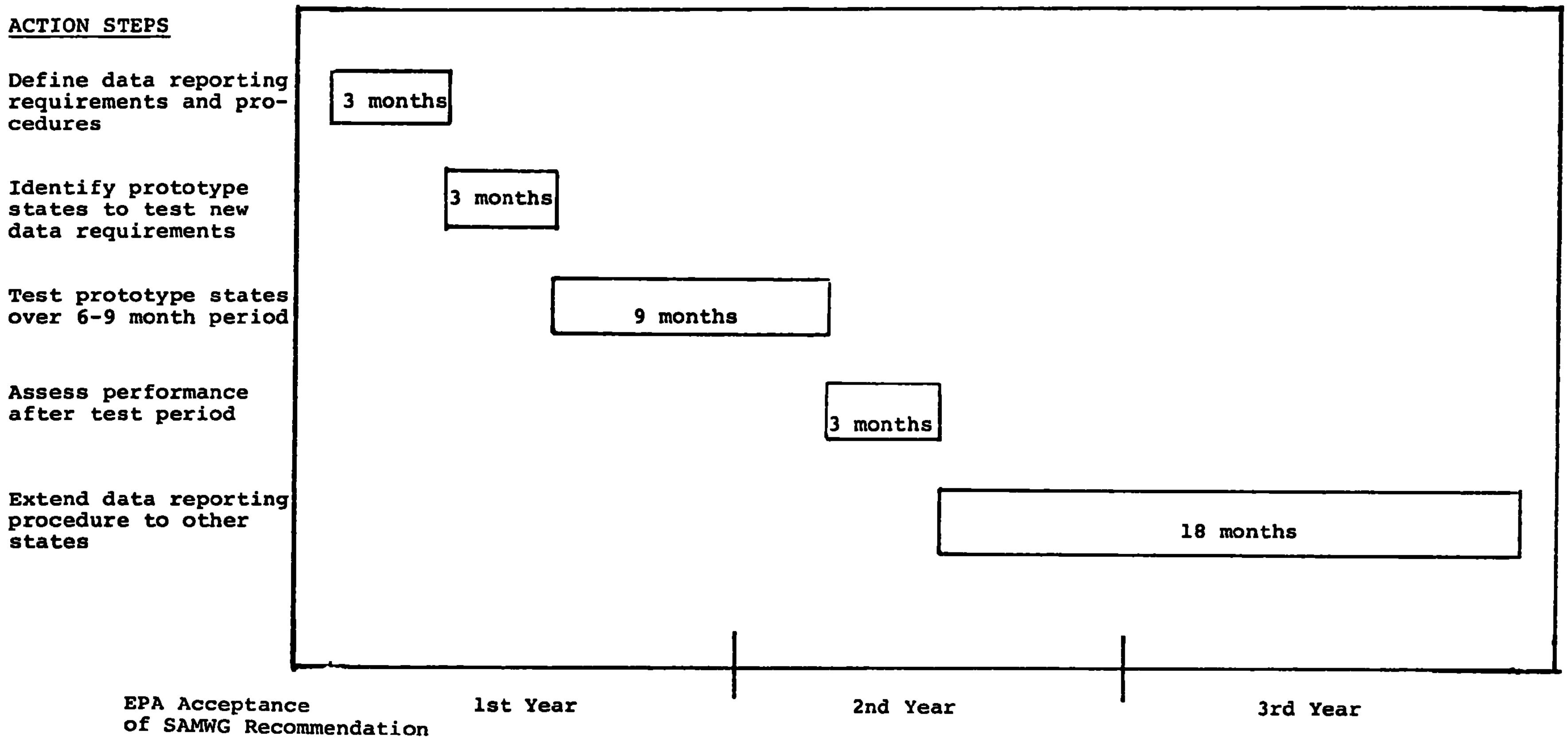
- The prototype states should test the new reporting procedures over a 6-9 month period.

- EPA should assess the performance of the new reporting procedures and revise these procedures accordingly.

- EPA should extend the new reporting procedures to the other states over an eighteen month period.

Exhibit 1-5, on the following page, indicates the relative phasing of these actions. Most of this work could be performed by EPA personnel. However, contractor assistance would be required to design and implement a management summary information system for use by EPA. Contractor assistance could be used also to develop detailed data reporting requirements consistent with the SAMWG recommendations. Overall contractor costs for these services would be approximately $75,000. Additional contractor support for CDHS installation and EPA personnel resource levels to coordinate these action steps would not exceed currently budgeted resource levels.

Exhibit 1-5
U.S. Environmental Protection Agency
Time-Phased Action Steps for Implementing the Distributed Air Quality System



(3) <u>EPA Should Refrain From Initiating Major New Changes to NEDS and SAROAD Until Currently-contracted Enhancements Have Been Completed and Tested.</u>

EPA should be able to achieve substantial reductions in current air data processing costs as a result of current design changes in its major systems. However, in order to insure that anticipated benefits are realized, EPA should refrain from making other major design changes to NEDS and SAROAD until Agency personnel have had an opportunity to confirm the expected efficiencies. In addition, these design changes should impact dramatically the accessibility of NEDS and SAROAD to field users and should provide more reliable edit, update and reporting services. Accordingly, the patterns of report retrievals and data submissions likely will change so that new processing statistics will be needed to determine whether additional opportunities for system efficiencies will be available.

However, minor changes in NEDS edit and update procedures should be incorporated into current enhancement efforts. In particular, current changes in the NEDS edit should be expanded to allow regions to run validation checks and to edit input cards against the NEDS user file. These changes will reduce the time lags associated with the correction of NEDS data and will cost approximately $10,000 in additional contractor funds.

Minor additional savings could be achieved through additional design changes depending upon the operating performance of the resulting air data systems. Six specific changes could further reduce operating costs and improve service, as follows:

- Modify the method for calculating SAROAD running averages
- Revise the timing of SAROAD summary and frequency of file creation
- Develop a SAROAD violations file
- Develop an enhanced NEDS regional edit
- Enhance current CDHS data edit and validation capabilities
- Develop better interfaces between state and federal emissions systems

More substantial design changes, as proposed by EPA program personnel, have not been found to reduce ADP costs. These proposed changes have included three areas, as follows:

- A change in the current design approach for the management of emission data would not reduce ADP expenditures.
- Consolidating NEDS with CDS would not reduce ADP expenditures.
- The transfer of air data systems from the UNIVAC to other EPA computers is not cost-justifiable at this time.

Overall, the currently planned technical changes combined with a longer-term commitment to a distributed ambient data system will reduce ADP costs without disrupting program activities.

(4) <u>EPA Should Conduct Immediately a System Audit of the Energy Data System (EDS) Before Assigning Additional Resources for Support and Development.</u>

EPA should conduct immediately a system audit of EDS to determine the value of the current system and to ascertain whether additional resources should be applied to support it. The statement of work for this audit should include five areas for investigation, as follows:

- The air program information requirements to be supported by the system
- The adequacy of the current system in satisfying these requirements
- The efficiency of the current system design
- Operating cost analyses in terms of the decisions for which the system provides information
- System management responsibilities

The estimated cost for the EDS audit is $15,000 and the elapsed time is estimated to be three months.

(5) <u>EPA Should Continue Its Support of CDHS and Insure Sufficient Resource Commitments.</u>

EPA should continue its support for CDHS by taking five system-related actions, as follows:

- Complete current identified modifications to CDHS
- Conduct a review of states to identify additional CDHS enhancements necessary to meet state information requirements
- Provide additional enhancements to CDHS reporting capabilities as specified by state users
- Improve current data handling procedures for CDHS by incorporating more extensive edit capabilities
- Review and develop interface procedures between state-operated systems and EPA-maintained systems

These actions will require an estimated $250,000 in contractor funds over a 12-month period. These contractor funds have been requested and granted by OAQPS and do not represent an additional level of effort above current budget requests.

(6) <u>MIDSD Should Conduct a Detailed Technical and Operating Analysis of the UNIVAC and Its Management.</u>

MIDSD should conduct a detailed technical analysis of the UNIVAC in four areas, as follows:

- Hardware malfunctions
- Throughput volume
- Channel capacities
- Operating system performance

This study should include an analysis of NCC processing characteristics and performance with other comparable UNIVAC installations. In addition, this study should analyze the adequacy of current staffing levels, management procedures, use of contractors and organization. As a result of this study MIDSD should develop an organizational and operations plan to reduce the overall failure rate at NCC. This analysis is estimated to cost $50,000 and could be completed within four months.

(7) <u>EPA Should Consolidate Its Air Data Processing Responsibilities and Formalize Its ADP Management Procedures.</u>

Air data system development and operation should be centralized within OAQPS and personnel with experience in managing development projects should be allocated by senior OAQPS management in proportion to the relative priorities of each division's needs. In addition, support for CDHS should be formalized as a high priority responsibility of OAQPS data systems managers and suitable training and assistance should be provided to regional coordinators to insure that states receive adequate data systems support.

MIDSD should provide assistance to OAQPS and other air data systems users to insure that good management principles are adhered to in the design, development and operation of ADP systems. This assistance should include four areas, as follows:

- Formalize statements of work to contractors at RTP to insure that feasibility studies are performed separately from system development
- Confirm that actual design and development activities are consistent with cost justifications as produced in feasibility studies
- Require formal ADP budgets from all NCC users and explain to program managers that program funds cannot always be interchanged with ADP sub-allowances
- Provide guidelines and accounting tools to NCC users to help them to develop improved methods for reducing processing costs

All these changes can be implemented internally by EPA staff without contractor assistance.

(8) <u>ADP Expenditures in Support of the Air Program Should Remain Relatively Stable Over the Next Two Years.</u>

Over the next two years ADP expenditures will remain relatively constant while contractor costs will decrease by an average of $250,000 annually.

Exhibit 1-6, on the following page, compares FY '76 estimated ADP expenditures by organization with the projected future annual costs. OAQPS will experience the most significant reduction in computer operating costs of approximately 37% compared to FY '76 expenditure levels. On the other hand, regional and state ADP costs will increase by approximately $275,000 as additional CDHS systems are installed and as local and regional personnel assume more responsibility for data analysis and modeling activities. Index has not been able to obtain projected ADP operating costs from either ORD Headquarters or laboratory personnel. However, we have estimated ORD operating costs to increase by 20% annually to reflect new system development and continued research analysis. This 20% assumed increase reflects growth in annual ORD expenditures of almost $350,000 over current levels. Exhibit 1-7, on the page following Exhibit 1-6, compares FY'76 estimated expenditures by system with projected future annual costs. The most significant reduction in operating costs will result from current technical SAROAD modifications. These technical changes plus the shift towards a hybrid approach for managing air quality data will reduce SAROAD computer costs by approximately 67%. On the other hand, CDHS costs will increase significantly as additional installations occur. EPA system managers have projected

Exhibit 1-6
U.S. Environmental Protection Agency
Projected ADP Operating Costs by
Organization

	1976 ($000)	1977 ($000)	1978 ($000)
OAQPS	1656	1360	1066
ORD	857	1000	1200
OE	54	50	50
Regions	300	300	400
States	300	380	475
TOTAL	3126	3090	3190

Exhibit 1-7
U.S. Environmental Protection Agency
Projected ADP Costs by System

	1976 ($000)	1977 ($000)	1978 ($000)
SAROAD	903	480	300
NEDS	150	180	180
CDS	105	100	100
RAPS	250	300	360
CHESS/CHAMPS	347	410	490
EDS	183	300	186
CDHS	300	381	475
NASN	54	54	54
Other OAQPS Systems	380	400	400
Other ORD Systems	200	230	290
Other (including unallocated regional usage)	254	255	355
TOTAL	3126	3090	3190

EDS operating costs to increase to $300,000 next year. However, we anticipate that an EDS system audit will identify areas where operating costs could be reduced back to current year levels. Personnel levels in support of the air data systems are expected to remain stable during the next two years at Headquarters, in the regions and in the states.

5. REPORT ORGANIZATION

The remainder of this report is divided into seven chapters and three appendices. Chapters II through V present findings regarding program strategy, user needs, data system functions and related problem areas. Chapter VI presents an evaluation of alternative actions to remedy identified problems and Chapter VII presents the resulting recommendations. Appendix A contains technical information on the air data systems reviewed. Appendix B contains cost derivation information as employed in the study analysis and Appendix C summarizes the results of a questionnaire completed by system users throughout the country.

II. AIR PROGRAM STRATEGY

II. AIR PROGRAM STRATEGY

This chapter presents an overview of the EPA air pollution control program and its associated strategies. The chapter is divided into six sections, as follows:

- Legislative Basis for the Air Program
- Regulations and Policy Development
- Stationary Source Air Program
- Mobile Source Air Program
- Research and Development, Monitoring and Enforcement Programs
- Future Developments

1. THE CLEAN AIR ACT AMENDMENTS OF 1970 PROVIDE THE LEGISLATIVE BASIS BY WHICH EPA HAS RESPONSIBILITY FOR DEVELOPING AND MANAGING PROGRAMS FOR THE REDUCTION OF NATIONAL AIR POLLUTION LEVELS.

With the enactment of the 1970 amendments to the Clean Air Act and subsequent formation of the Environmental Protection Agency (EPA), the role of the federal government changed from one of limited authority in the abatement of air pollution to one of nationwide involvement in air pollution control. In 1955, the Department of Health, Education, and Welfare was authorized to conduct research and provide technical assistance to state and local communities in their attempt to prevent and control air pollution. In legislation of 1963 and 1967, Congress reaffirmed the policy of state and local responsibility for air pollution control while expanding the federal government's research and development activities. The Clean Air Act Amendments of 1970 provided the basis for federal leadership in the development of national, regional, state and local programs to prevent and control air pollution.

(1) In Accordance With the Act, EPA Has Set National Primary and Secondary Ambient Air Quality Standards for Each Criteria Pollutant.

In 1970, EPA published criteria documents which described the current scientific knowledge of the effects on public health and welfare which could be expected from the presence of various pollutants in ambient air. Based on these documents, in April 1971 EPA established primary and secondary standards for six major pollutants, as follows:

- Sulfur dioxide (SO_2)
- Particulate matter
- Carbon monoxide (CO)
- Hydrocarbons (HC)
- Nitrogen oxide (NO_x)
- Photochemical Oxidants

Exhibit 2-1 on the following page, summarizes the major source of each pollutant and the associated general health-related area of concern.

National primary ambient air quality standards define levels of air quality necessary to protect the public health with an adequate margin of safety, while secondary standards are designed to protect the public welfare, specifically property, vegetation, and wildlife.

As published in the Federal Register, each ambient air standard specifies a reference method for sampling and analyzing the ambient air and the reference conditions to which measurements are to be corrected. Under the Act states may establish ambient air quality standards more stringent than the national standard to prevent significant deterioration of existing air quality within their state.

(2) <u>States Are Required by Law to Develop and Carry Out State Implementation Plans (SIP's) to Achieve and Maintain National Ambient Air Quality Standards (NAAQS) Established by the EPA.</u>

Exhibit 2-1
U.S. Environmental Protection Agency
Summary of Criteria Pollutants and
Associated Sources and Areas of
Concern

POLLUTANT	MAJOR SOURCE	GENERAL AREA OF CONCERN
Sulfur Dioxide	Stationary Sources	Respiratory Illness
Particulate Matter	Stationary Sources	Respiratory Illness
Carbon Monoxide	Mobile Source	Lethal Gas
Hydrocarbons	Mobile Sources	Contributes to Formation of Oxidants
Nitrogen Oxides	Mobile Sources	Respiratory, Contributes to Formation of Oxidants
Photochemical Oxidants	Formed in the Atmosphere from NO_x and HC	Respiratory Illness

The ambient air standards became the basis for the development of SIP's in 1972. State plans specified the steps to be taken to control emissions from existing sources of pollutants in order to achieve ambient air quality standards. The Clean Air Act gives EPA the authority to approve or disapprove SIP's and all subsequent revisions that the states may develop. Furthermore, EPA was required to develop the implementation plan for any state if one of the following three conditions existed:

- The state failed to submit a plan within the time prescribed by the Clean Air Act
- EPA determined that the plan did not meet the requirements of Section 110 of the Clean Air Act
- The state failed to make appropriate revisions as deemed necessary by the EPA to achieve NAAQS

The primary responsibility for enforcement of the SIP rests with the state while EPA has concurrent authority to enforce the SIP's, when necessary.

(3) Although the Clean Air Act Emphasizes State Responsibility, the Clean Air Act Gives EPA Special Authority in Certain Areas That Are Interstate in Nature or Pose a Severe Threat to Health

Special federal authority exists to control new stationary sources of criteria pollutant emissions, motor vehicle emissions, sale of fuel and fuel additives, and emission of hazardous air pollutants.

EPA directly regulates new stationary sources by setting uniform national standards for various new sources of pollution. The standards of performance published in the Federal Register are designed to limit emissions through the use of the Best Available Control Technology (BACT).

Since the abatement of mobile source pollution requires national standards, the Clean Air Act dictated a schedule for the abatement of automobile pollution. In accordance with the Act, EPA prohibits the sale of new cars unless the cars are based upon a tested and approved prototype model.

In addition, EPA, in coordination with the Federal Energy Administration (FEA), has written regulations concerning fuels, and fuel additives. In 1974, EPA and FEA coordinated actions to make non-leaded gasoline available at service stations so as not to interfere with the performance of the catalyst pollution control devices.

After the 1973 oil embargo, Congress passed the Energy Supply and Environmental Coordination Act (ESECA) of 1974. ESECA required the FEA to issue orders prohibiting power plants with coal burning capabilities from burning natural gas or oil as a primary source of energy. The act gives EPA responsibility for the certification of FEA prohibition orders. Before an FEA order becomes effective EPA must determine if primary air quality standards will be violated as a result of a conversion. The Clean Air Act was amended in 1975 to authorize EPA to suspend temporarily fuel and emission limitations for those sources which have been issued orders by FEA and certified by EPA.

EPA has set national standards for air pollutants that have proven to contribute to increased death rates. Once promulgated, these standards are applicable to all new and existing sources.

(4) <u>As Specified in the Clean Air Act, EPA Has Established a National Research and Development Program for the Prevention and Control of Air Pollution</u>.

The Office of Research and Development (ORD) conducts a comprehensive and integrated research and development program in support of the objectives of the Clean Air

Act. ORD supports the Air Program in four general areas, as follows:

- Providing research on both the short and long-term effect of air pollutants on human health to be used as the scientific basis for standards and regulations
- Prescribing standard methods to measure and assure quality control in programs to assess environmental quality, implement plans and enforce strategies
- Identifying cost-effective pollution control technology which is capable of achieving short-term abatement objectives and possesses long-term viability
- Providing expert scientific assistance through the dissemination of technical information and technological transfer of capabilities

2. <u>THE AIR PROGRAM STRATEGY IS DESIGNED TO ACHIEVE AND MAINTAIN NATIONAL AMBIENT AIR QUALITY STANDARDS</u>.

EPA has developed an air program strategy to achieve the objectives of the Clean Air Act by taking six actions, as follows:

- Divide the nation into Air Quality Control Regions
- Encourage states to develop and enforce implementation plans to achieve and maintain NAAQS
- Develop performance standards for new stationary sources
- Conduct programs to prevent the further deterioration of air quality

- Develop a national mobile source air program
- Establish national emissions standards for hazardous air pollutants

These actions are designed to specify local air pollution abatement objectives and to achieve these objectives through the control of existing and new sources of pollution.

(1) <u>EPA has Divided the Nation Into Air Quality Control Regions (AQCR's) in Order to Develop and Implement Plans Which Specify the Manner in Which NAAQS Will Be Met and Maintained.</u>

Under Section 107 of the Clean Air Act each state has the primary responsibility for assuring air quality within its boundaries. The country is divided into 247 AQCR's which reflect the unique features of local air pollution features. An AQCR is a geographic unit within which states analyze existing ambient air conditions, develop plans, and determine priorities for action. Each region is classified separately with respect to each criteria pollutant so that the time and resources expended in developing the SIP can be commensurate with the complexity of the air pollution problem. Each region is classified into one of three categories, as follows:

- Priority I - most restrictive
- Priority II
- Priority III - least restrictive

Exhibit 2-2, on the following page, summarizes the classification mechanism for each pollutant.

(2) EPA's Air Strategy Requires States to Develop and Implement Plans to Achieve and Maintain NAAQS for Each AQCR Within Their Geographic Jurisdiction.

State Implementation Plans (SIP) detail the steps to be taken to abate and control emissions from existing sources within an AQCR to insure attainment of air quality standards. An implementation plan includes four major elements, as follows:

- Information on existing air quality and source emissions
- Control plan for achieving standards
- Programs for monitoring ambient air quality and source emissions
- Description of legal authority

EPA may require a SIP revision under any of three circumstances, as follows:

Exhibit 2-2
U.S. Environmental Protection Agency
Summary of Classification of Regions

POLLUTANT	BASIS FOR CLASSIFICATION	APPLICABLE CLASSES	METHOD OF CALCULATION
Sulfur Oxides	. Measured Ambient Air Quality, or . Estimated Air Quality in Area of Maximum Concentration	I,II or III	Annual Arithmetic Mean, 24-hour Maximum and 3-hour Maximum
Particulate Matter	. Measured Ambient Air Quality, or . Estimated Air Quality in Area of Maximum Concentration	I,II or III	Annual Geometric Mean and 24-hour Maximum
Carbon Monoxide	. Measured Ambient Air Quality, or . 1970 Bureau of Census Population Information	I or III	1-hour Maximum and 8-hour Maximum
Nitrogen Dioxide	. Measured Ambient Air Quality, or . 1970 Bureau of Census Population Information	I or III	Annual Arithmetic Mean
Photochemical Oxidants and Hydrocarbons	. Measured Ambient Air Quality, or . 1970 Bureau of Census Population Information	I or III	1-hour Maximum

- Changes of national primary or secondary national air quality standards
- Evidence to EPA that the SIP is substantially inadequate to achieve NAAQS
- Changes in the availability of methods and control technology necessary for source compliance

In addition, in the development of a SIP, the state must consider the effect of urban and economic growth on a region's future air quality and adapt its control programs accordingly.

(3) New Source Reviews (NSR) and Environmental Impact Statements (EIS) Are Mechanisms to Prevent the Further Deterioration of Air Quality.

As a result of a 1972 District Court decision and subsequent affirmation in 1974 by the U.S. Supreme Court, EPA is required to disapprove SIP's which do not provide a plan to prevent the significant deterioration of air quality in areas that did not exceed national ambient standards during 1974 for sulfur dioxide and particulate matter. Accordingly, EPA published regulations in the Federal Register in 1974 establishing three classes of AQCR's based on how much increase of ambient concentrations of sulfur dioxide and particulate matter could be permitted without significant deterioration of air quality. The three classes are defined, as follows:

- CLASS I — AQCR's where practically any air quality deterioration would be considered significant
- CLASS II — AQCR's where well-planned growth would not be considered to add significantly to the air quality deterioration
- CLASS III — AQCR's where intense industrial growth is desired but national air standards for particulate and sulfur dioxide must be maintained

EPA designated all ACQR's as Class II, originally; however, states may request redesignation of a region in order to meet various social and economic objectives. For purpose of significant deterioration determination, areas designated as Class I or II are limited to specific increases in particulate matter and sulfur dioxide concentrations above January, 1975 ambient air quality levels for those pollutants.

Nineteen source categories have been established for the review of new or modified sources commencing construction or modification after June, 1975. The non-significant deterioration reviews are designed to determine the effect on air quality of the proposed new source or source modification in conjunction with changes in emissions from other sources in the area of concern. EPA may delegate the responsibility for implementing procedures for conducting non-significant deterioration reviews to the individual states.

The National Environmental Policy Act of 1970 requires each federal agency to prepare a statement of environmental impact in advance of each major action that may affect ambient air quality significantly. EIS's provide a mechanism by which federal agencies, the public, and Congress can review the air quality consequences associated with a pending government decision. The Council on Environmental Quality (CEQ) is responsible for coordinating actions relating to EIS production.

(4) <u>EPA's Mobile Source Program Is Designed to Reduce the Levels of Harmful Pollutants Emitted From Motor Vehicles.</u>

In 1971, motor vehicles contributed 50% of all carbon monoxide (CO), 40% of all hydrocarbons (HC) and 25% of all nitrogen oxides (NO_x) emitted into the air within the United States. The Clear Air Act legislated strict mobile source emission standards for the control of these three pollutants. Exhibit 2-3, on the following page, summarizes the mobile source abatement goals as legislated in 1970. In 1973, EPA granted a one-year extension for the 1975 statutory standards.

The choice of how to meet the legislated standards has been left to mobile source manufacturers. In the early 1970's, most U.S. automakers elected to use catalysts as a means of converting HC and CO exhaust emissions into carbon dioxide and water. The catalyst represented a control technology which could reduce emissions to required levels while protecting auto manufacturers' capital and technical investment in the internal combustion engine.

(5) EPA Established National Emissions Standards for Hazardous Air Pollutants (NESHAP).

EPA has established emission standards for air pollutants which have been found to be harmful to the public health and to which no ambient air quality standard is applicable. Three hazardous pollutants have been identified, as follows:

- Asbestos
- Beryllium
- Mercury

Exhibit 2-3
U.S. Environmental Protection Agency
Summary of Legislated Mobile Source
Pollution Abatement Goals

POLLUTANT	GOAL
Hydrocarbons	1975 new models reduction of 90% of 1970 model emissions
Carbon Monoxide	1975 new models reduction of 90% of 1970 model emissions
Nitrogen Oxide	1976 new models reduction of 90% of 1971 model emissions

Once promulgated, NESHAP's are applicable to both new and existing stationary sources. States may submit procedures for implementing and enforcing NESHAP's for stationary sources located within their geographic jurisdiction. Upon review and approval of the procedures EPA may delegate the authority to implement and enforce the NESHAP.

3. <u>EPA'S STATIONARY SOURCE AIR STRATEGY IS DESIGNED TO CONTROL THE LEVELS OF EMISSIONS EMANATING FROM MAJOR POINT SOURCES.</u>

The identification of point sources is the first step in attaining NAAQS. A point source is defined as any facility capable of emitting more than 100 tons per year of any single pollutant without pollution controls. Source parameters and emissions information are needed to develop and evaluate control strategies to be used in setting compliance schedules for sources in non-attainment areas. Of 20,000 major stationary sources of air pollution in 1975, 84% are complying currently with emissions regulations or are meeting abatement schedules.

(1) <u>SIP's Detail the Steps to be Taken for the Abatement and Control of Emissions for Existing Point Sources.</u>

As a part of SIP development, states identify major point sources, set source emissions limitations, and establish schedules for compliance so that total emissions are reduced to levels which permit ambient air quality standards to be achieved. Included in the SIP are provisions for monitoring and gathering data to track source compliance with state-established timetables.

(2) Emission Control Regulations (ECR) Are Based on Control Strategies That Are Capable of Achieving Desired Air Quality While Minimizing Overall Costs.

Emission control regulations are legally enforceable emission control requirements which consist of emission standards covering significant source types. The development of the ECR's is based on a review of four areas, as follows:

- Existing ambient pollution levels
- Source and emissions inventories
- Available control technology
- Projections of increased source emission due to urban growth

In the formulation of control regulations, states evaluate various strategies in terms of their capability of achieving desired air quality in the least costly or disruptive manner.

(3) EPA Has Established New Source Performance Standards (NSPS) to Control Emissions From New Stationary Sources.

EPA has promulgated NSPS requiring all new sources of pollution within specified industries to install the best available control technology as an effective means of controlling current and future air pollution levels. Stationary sources are classified by types and sizes for the formulation of NSPS. To date, EPA has established standards for twenty industrial activities, which can be grouped in eleven categories, as follows:

- Fossil Fuel-Fixed Steam Generators
- Municipal Incinerators
- Cement and Asphalt Plants
- Sulfuric and Nitric Acid Plants
- Sewage Treatment Plants
- Petroleum Refineries
- Phosphate Fertilizer Industry
- Iron and Steel Plants
- Zinc, Lead and Copper Smelters
- Coal Preparation Facilities
- Aluminum Reduction Plants

States may submit procedures for implementing and enforcing NSPS for sources within their geographic jurisdiction. Upon review and approval of the procedures EPA may delegate to the state the authority to implement and enforce the NSPS.

(4) Models and Monitoring Are Used to Evaluate Control Strategies and to Revise SIP's Where Necessary.

EPA and states have utilized models as tools to assist in the development of programs that lead to the attainment of air quality standards. Monitoring information and diffusion models are used to ascertain the effect that the distribution of emissions from a point source has on the ambient air quality. Monitoring, itself, is expensive in terms of manpower and resources; therefore, models are used as supplements to the monitoring effort. In addition, accurate models provide a more comprehensive picture of ambient air quality than do monitors, which tend to reflect air quality only in the immediate vicinity of the individual stations.

Once a control strategy is selected, state offices employ air quality sampling to monitor progress toward achieving SIP goals and national standards. Monitoring also serves as a validation tool by which model developers observe how well the relationship between air quality and source emissions is simulated. In this way, modeling and monitoring supplement each other in identifying appropriate control strategies.

4. <u>EPA'S MOBILE SOURCE AIR STRATEGY IS DESIGNED TO LIMIT MOBILE SOURCE EMISSIONS BY REGULATING MANUFACTURERS, REGULATING MOTOR FUEL DISTRIBUTION AND FACILITATING THE DEVELOPMENT OF TRANSPORTATION PLANS.</u>

Although the Clean Air Act legislated strict emissions standards, the proliferation of automobiles has hindered the abatement process and contributed to a national energy shortage. In 1970, mobile vehicles traveled over a trillion miles. Statistics have revealed that the majority of these auto trips were less than five miles long and that 85% of commuting was done by car with usually just the driver inside. This analysis demonstrated that the mobile source problem had three facets, as follows:

- Development of effective means of reducing pollutants emanating from motor vehicles
- Development of efficient transportation systems
- Land use planning and other methods of reducing geographical concentrations of automobile pollutants

Consequently, solutions to transportation and land use issues are needed also to improve air quality and energy resource utilization.

(1) <u>Mobile Source Program Personnel Interact With Automobile Manufacturers and Importers to Certify Compliance With Emissions Standards as Legislated in the Clean Air Act.</u>

Each year motor vehicle manufacturers apply to EPA for certification of compliance with emission standards. EPA provides this certification by testing prototype vehicles prior to their introduction to assure that each new model is engineered to conform to emissions regulations over its useful life. As a part of this program, EPA is authorized to inspect records, procedures, and production facilities.

All motor vehicles manufactured abroad for sale in the United States also must meet EPA auto emission standards. Foreign manufacturers have been given the same compliance schedules as domestic producers. EPA coordinates its certification activities with the U.S. Customs Department for the importation of foreign-made vehicles. Any individual who purchases a car overseas for importation into this country is responsible for verifying compliance with emission standards.

(2) <u>EPA Is Working With States and Municipalities to Establish Transportation Control Plans (TCP's), Inspection Maintenance Programs and Vapor Recovery Programs.</u>

The automobile emissions problem is especially serious in certain urban areas; consequently, EPA has initiated transportation control strategies to deal with the problems unique to these areas.

Transportation Control Plans are designed to reduce aggregate emission of pollutants from vehicles by decreasing total vehicle miles traveled (VMT) through five strategies, as follows:

- Limited access areas
- Mass transit improvements
- Traffic flow improvements
- Carpool programs
- Priority bus lanes

TCP's impose a social burden on the general public because they affect the way people move from one place to another and thereby impact urban lifestyles.

In addition, EPA is providing program guidance and technical support to states and localities in developing emissions inspection maintenance programs. These programs are designed to insure that vehicles continue to achieve the same level of emission control as when they were manufactured. EPA also conducts a vapor recovery program to reduce the emission of hydrocarbon vapors at gasoline filling stations as fuel is transferred from storage tanks and trucks to stations and from gasoline pumps to vehicles.

(3) The Mobile Source Program Monitors the In-Use Performance of Motor Vehicles to Determine Emission Levels After the Vehicles Have Been Purchased and Driven by Their Owners.

Emission performance tests of vehicles in actual use provide data to assure that vehicles continue to fulfill federal emission requirements. Where necessary, enforcement orders are given to manufacturers stipulating a recall of vehicles for corrective action. Under the Clean Air Act, car manufacturers must provide a warranty of five years or 50,000 miles on the emission control system of their cars.

The Office of Research and Development (ORD) derives its activity priorities from the objectives set by EPA in fulfillment of the legislative mandate of the Clean Air Act. Exhibit 2-4, on the following page, summarizes the ORD activities in support of the air program.

Air pollution generally has been recognized as an environmental influence on the public health. Consequently, ORD conducts on-going research directed toward the understanding of the relationship between air pollution levels and public health. Research continues on the six pollutants for which criteria documents have been published; however, continual emphasis is given also to pollutants not presently controlled. In support of this research, EPA conducts toxicological, clinical, and epidemiological studies to gather information on four areas of concern, as follows:

- Exposure/response of various pollutants over time
- Development of standards that provide an adequate margin of safety
- Health benefits in meeting standards
- Health risks in exceeding standards

Exhibit 2-4
U.S. Environmental Protection Agency
Summary of Research and Development
Activities in Support of Air
Program

AREA OF ACTIVITY	ASSOCIATED OBJECTIVES
Health Effects	• Understand relationship between air quality and public health • Identify pollutants to be controlled • Develop criteria documents to be used for standard setting
Technological Assessment	• Identify control technologies and ascertain effectiveness viability, and economic implications • Evaluate health implications of control equipment and future technological processes
Scientific Methodologies	• Understand atmospheric processes governing transport and dispersion of pollutants from various sources • Develop, evaluate, and validate models for predicting and analyzing emissions from a variety of sources
Quality Assurance	• Develop standardized measurement methods, reference methods, and manuals • Evaluate laboratory instruments and data handling capabilities

Health effects research includes exposure/response relationships of both criteria and non-criteria pollutants by themselves and in combination with each other.

ORD provides emission characteristics, measurement methods, and health-related data for new control devices, fuel additives and technical processes. For example, the adoption of catalytic converters to control pollutants from mobile sources is under review by research scientists. Toxicological studies are being conducted to establish the health implications of sulfates and sulfuric acid mist that are produced by catalyst-equipped cars. Similarly, as new source performance standards are proposed, ORD identifies and evaluates available technology to meet the standards in the most cost-effective manner.

In addition, ORD develops air quality simulation models for predicting ambient concentration levels and for analyzing emissions from various sources. In order to develop effective models, scientific knowledge on the transport and dispersion of pollutants in the atmosphere is needed. During the past three years, ORD has conducted a major Regional Air Pollution Study (RAPS) in St. Louis to develop, evaluate, and validate urban air quality models for the criteria air pollutants.

Air pollution monitoring methods require high performance reliability at very low concentration levels because of the nature and degree of ambient pollution levels. Although ORD does not have responsibility for air monitoring, it is responsible for developing and implementing a comprehensive quality assurance program. The quality assurance program consists of three activities, as follows:

- Develop and publish standard measurement methods and quality control manuals
- Provide technical assistance to state and regional offices for establishing quality control programs
- Review state proposals for assuring quality control

As necessary, performance audits of monitoring sites are conducted to assure that accurate and defensible data are being gathered.

(2) <u>EPA's Primary Ambient Air Monitoring Goal Is to Provide Data of Adequate Quality to Arrive at Decisions Regarding the Attainment of NAAQS.</u>

Data produced from the ambient air monitoring networks form the basis for the evaluation of air quality trends and for the evaluation of the success of a SIP. Moreover, monitoring data provide information to evaluate which areas are attaining NAAQS and which will require SIP revisions in order to attain air quality standards.

EPA is concerned especially with three aspects of monitoring data, as follows:

- Enhancement of the quality of data through development of site and laboratory quality assurance programs
- Optimal collection and use of data in order to minimize collection of unnecessary data
- Review of existing data for their adequacy in support of SIP evaluation and strategy development

The Standing Air Monitoring Work Group (SAMWG) is an EPA task force analyzing the continuing need for and use of monitoring data at Headquarters, the regions, and in state and local agencies. Preliminary findings of SAMWG indicate that less monitoring data will be needed in the future by EPA to assess national trends, but that more short-term monitoring will be needed by state and local agencies to pursue abatement and enforcement activities.

(3) <u>The Air Enforcement Program Is Designed to Identify and Investigate Suspected Violations and Establish Compliance Schedules to Bring Violators into Conformity With Published Regulations.</u>

EPS's current air enforcement programs consist of nine major activities, as follows:

- Enforcement of State Implementation Plan
- Regional and state coordination in monitoring compliance schedules of major stationary sources
- Enforcement of transportation control plans
- New Source Performance Standards enforcement
- NESHAPS enforcement
- Establishing priorities in non-attainment AQCR's for resource allocation
- Providing technical, legal, and case development assistance to state and regional programs
- Participation in ESECA coal conversion program
- Participation in non-significant deterioration programs

EPA efforts are aimed at encouraging states to assume a larger role in enforcement as a key to national compliance.

6. THE FUTURE ACTIVITIES OF THE AIR PROGRAM WILL BE AFFECTED BY THE IDENTIFICATION OF NEW POLLUTANTS, THE DEVELOPMENT OF NEW TOOLS TO MEET AIR PROGRAM OBJECTIVES AND THE CONTINUING DELEGATION OF AUTHORITY TO REGIONAL AND STATE OFFICES.

The current air program strategy and activities are involved primarily in the planning for and control of the criteria pollutants through the SIP and NSPS mechanisms. These activities will be affected over time as new pollutants are identified and as new methods are developed for controlling pollution. In addition, shifting emphasis from Headquarters to the field likewise will change the levels and types of activities in which pollution control officials will be involved.

(1) ORD Is Giving Special Emphasis to the Identification of Pollutants Which Are Not Controlled Presently and Which Would Require the Promulgation of Additional NAAQS and NESHAP's.

In the past, health effects research has been conducted primarily to ascertain the diseases which may be related to criteria pollutants. Although this activity will continue, ORD will focus also on the environmental factors which are associated with the causes of three categories of diseases, as follows:

- Chronic Respiratory Diseases
- Cardiovascular Diseases
- Cancer

Through this effort new pollutants may be identified which will require control. Emerging problems such as vinyl chloride, sulfates and other potential hazardous substances will receive increased attention. Similarly, since the majority of most people's time is spent indoors, EPA will expand its effort in assessing the health effects associated with indoor pollutant levels.

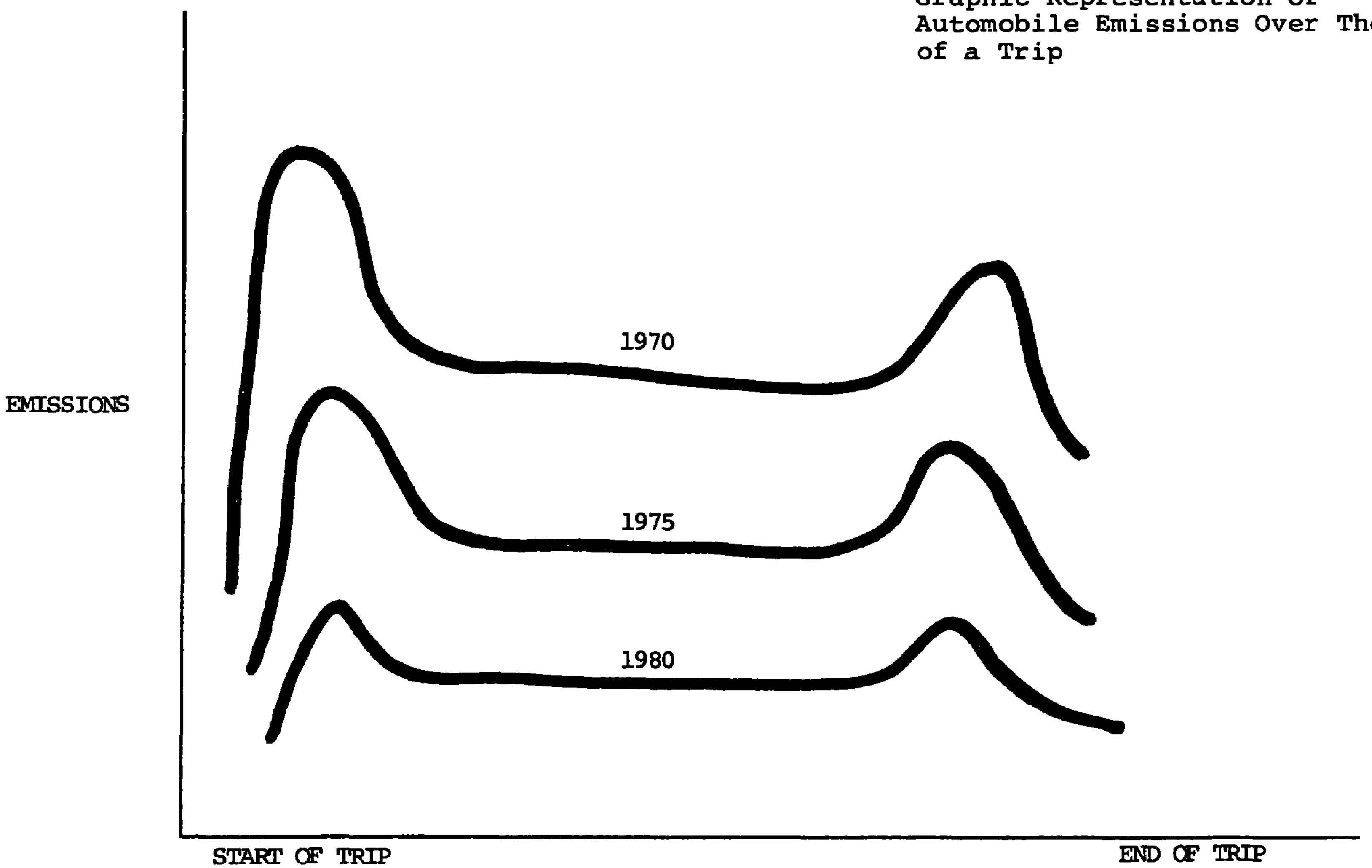
(2) <u>As the Air Program Evolves, New Approaches Will Be Developed to Assist in Accomplishing the Objectives of the Clean Air Act</u>.

As the SIP revision process continues, the rollback strategy basis for plans is proving to be inadequate and not appropriate in many cases. States increasingly are using simulation models to evaluate control strategies. Consequently, programs such as RAPS play a key role in developing models which can simulate accurately the dispersion of pollutants into the atmosphere. Emission inventories and demographic information are important input data to support these efforts.

As motor vehicle emissions continue to be reduced because of manufacturer design changes and the use of the catalytic converter, transportation control plans will shift emphasis from decreasing total vehicle miles traveled to reducing the total number of trips. Exhibit 2-5, on the following page, illustrates the pattern of emissions over the course of an automobile trip. The start and end of the trip represent the periods of largest concentration of emission. Various air program planners from the State of California believe that a strategy designed to reduce the number of trips, rather than the total vehicle miles, may prove effective in reducing aggregate emissions as part of a TCP.

Different approaches may be needed to reduce ambient air pollution levels because of unique characteristics of specific pollutants. For example, ozone and other oxidants have been considered to be an urban problem because they were produced by photochemical reactions of local emissions of hydrocarbons and nitrogen oxides. In recent years, however, standard violations of ambient oxidant levels have been observed in several rural areas of California and other states, far from major sources of hydrocarbons and nitrogen oxides. Accordingly, oxidant transport has been characterized as a phenomenon

Exhibit 2-5
U.S. Environmental Protection Agency
Graphic Representation of
Automobile Emissions Over The Course
of a Trip



which can affect population clusters distributed over relatively large geographic areas. The atmospheric formation of oxidants is a complex issue of which there is limited scientific understanding. Similarly, there is little knowledge of the atmospheric chemistry associated with sulfate formation and transport. As a result, fundamental questions arise concerning the chemicals, methods and siting of monitoring stations. Furthermore, these questions pose added difficulties with regard to the use of data to plan and formulate control strategies.

(3) <u>EPA Is Striving to Delegate Authority and Responsibility for Accomplishing the National Air Program Objectives to Regional and State Offices.</u>

The air program has evolved from a centralized group setting standards and national policy to one of regional and state development, implementation, and enforcement of programs to achieve and maintain national ambient air quality standards. While various centralized functions will continue, regional and state offices have the responsibility for the success or failure of accomplishing national air program objectives within their geographical areas. To this end, EPA is training personnel and allocating resources to regions in the form of increased regional budgets.

This effort is augmented by regions supplying grant monies to states in support of air pollution control programs. The Office of Air Quality Planning and Standards (OAQPS) has established the Air Pollution Training Institute and Air Pollution Technical Information Center to train state and local personnel and to distribute instruction manuals and various technical publications in support of air programs.

* * * * * * * * *

In summary, the air pollution control strategy represents a cooperative effort between the federal, state and local governments in reducing the level of ambient air pollution as specified by the Clean Air Act. The federal government develops air quality standards and finances research into related technical and medical areas; the state and local governments establish and implement plans to achieve the federal air quality standards. Most air quality control efforts, at present, are designed (1) to reduce emissions from current large point sources, (2) to control the introduction of new point sources of pollution, (3) to modify the design of motor vehicles in order to reduce emission of pollutants and (4) to influence driving patterns in order to reduce the total amount of mobile source pollutants generated.

As the air program continues to develop, increasing emphasis will be placed upon refining the methods of identifying and rectifying pollution problems. Improvements in modeling and control technology will provide pollution control experts with better tools to decrease the total amount of ambient air pollution while minimizing the associated social or economic disruption. In addition, as the program continues to mature, proportionately more technical and personnel resources may be required in the field to implement and enforce the control plans whereas proportionately fewer resources will be retained centrally for national program planning, development and technical assistance.

III. AIR DATA NEEDS

III. AIR DATA NEEDS

This chapter discusses user information requirements associated with air program activities. The chapter is divided into three sections, as follows:

- Information Needed to Plan and Review Air Program Strategies
- Information Needed to Implement Abatement Plans and Programs
- Information Needed to Analyze Problem Areas and to Develop New Programs

1. REGIONAL AND STATE OFFICES REQUIRE COMPREHENSIVE, ACCURATE AND CURRENT DATA TO REVISE STATE IMPLEMENTATION PLANS WHEREAS HEADQUARTERS NEEDS SELECTED DETAIL AND ANNUAL SUMMARY INFORMATION FOR VARIOUS PLANNING AND REVIEW FUNCTIONS.

Different types of information are needed at various levels of involvement in the air program to plan and review strategies in order to achieve the goals of the Clean Air Act. Exhibit 3-1, on the following page, summarizes the information need characteristics associated with the four major planning and review activities.

(1) EPA Headquarters Needs Summary Information to Develop and Rank National Air Program Plans and to Report to Congress on the Progress in Achieving the Objectives of the Clean Air Act.

Exhibit 3-1
U.S. Environmental Protection Agency
Planning and Review Information
Needs

	Ambient Air Data			Emissions Data						
	Accuracy	Currency	Completeness	Accuracy	Currency	Completeness	Compliance Data	Other Data	Summary or Detail	National or Local
Rank National Air Program Plans and Report to Congress		Annual	✓		Annual		Current	Technological, Research, Fuels Consumption, Demographic	S	N
SIP Development		Annual			Annual	✓			D	L
SIP Revision	✓	Quarterly		✓	Annual	✓	Current Accurate	Meteorological Demographic Topographic Technology Availability	D	L
TCP Development		Annual HC, CO, NO_x			Annual			Demographic Meteorological Traffic Flow Fuel Availability	D	L

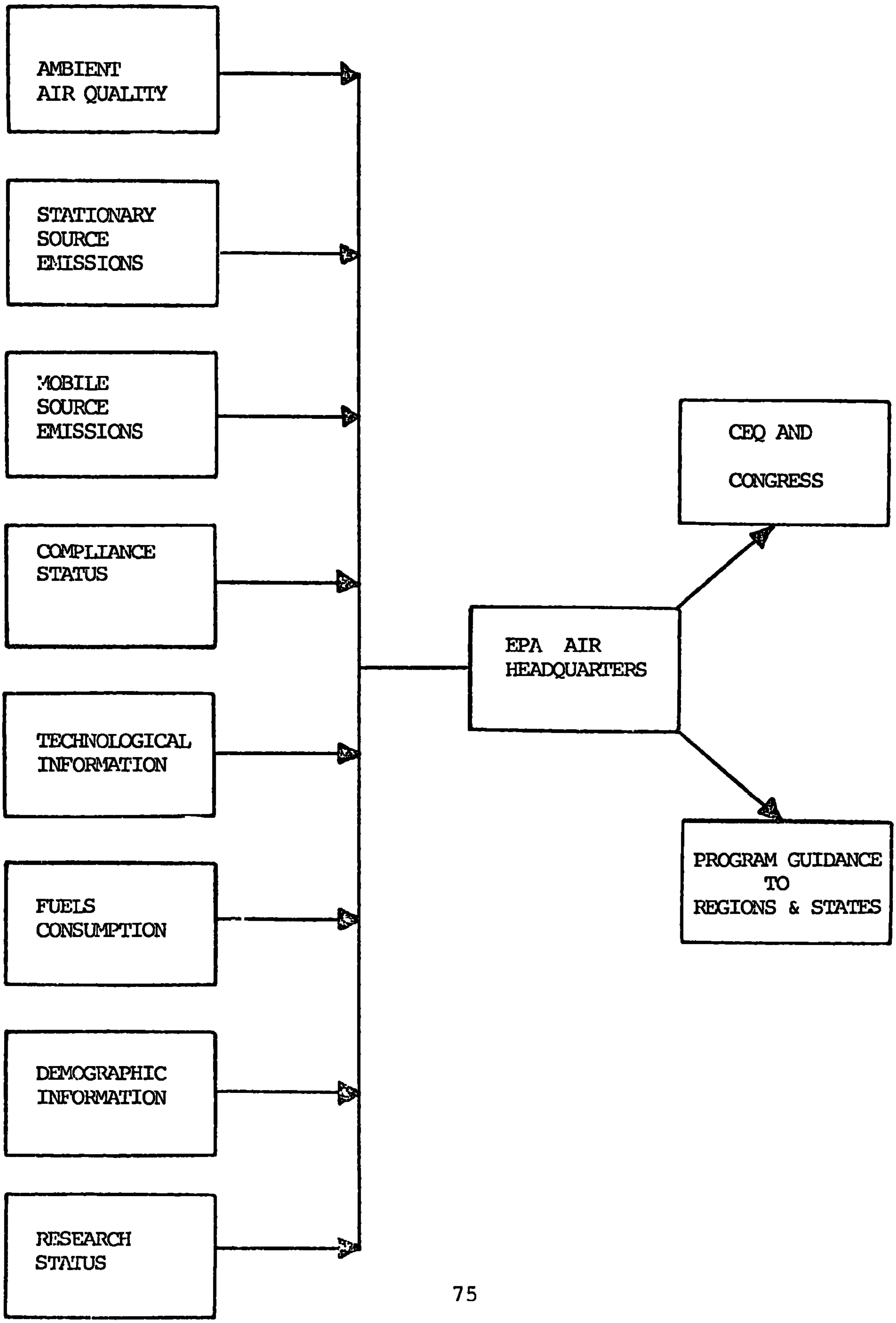
✓ = Important Characteristic

blank represents less important characteristic

EPA Headquarters uses summary information on the status of the national air program in order to report to the public on the condition of the environment and to prepare program guidance for regional and state offices. This information includes air quality, stationary source emissions and compliance status data, which are collected primarily from state and local agencies. In addition, EPA itself assembles mobile source emissions, technological, fuels and research information for use in program planning. Exhibit 3-2, on the following page, illustrates the flows and uses of this information.

Each fiscal year EPA prepares an Annual Air Program Planning Guidance document for regional, state and local offices to review in preparing program plans. The environmental priorities contained in the guidance are to assist regions and states in allocating resources to achieve and maintain national standards. EPA Headquarters uses trend information on ambient air quality together with knowledge of the availability of appropriate control technology to formulate and order program plans. In addition, information on the size and characteristics of

Exhibit 3-2
U.S. Environmental Protection Agency
Information Needed for EPA Air
Headquarters Planning and Reporting



population segments affected by current pollution levels and collected by other agencies is used to assess the scope of the national air pollution problem. EPA also uses air quality and emissions data summaries to determine national manpower and resource needs. Where possible, this summary information should be complete and no more than one year old.

The Council on Environmental Quality provides information about the condition of air quality and trends in air quality factors to environmental scientists, Congress and the general public. Annual government publications and reports provide information on six areas, as follows:

- Ambient Air Quality
- Stationary Source Emissions
- Mobile Source Emissions
- Compliance Status
- Research Status
- Technological Information

Other government agencies report annual national fuel consumption and demographic information.

EPA requires historical data to evaluate the progress of national control efforts. Since the 1970-71 period is the baseline period for evaluating overall progress toward attainment of NAAQS, EPA would like to have data representing ambient air quality for each period since 1970-71 to prepare trend reports and analyses. Trend information concerning emission reduction requires comparable historic data for comparison against current emissions levels. However, most state agencies had only limited air monitoring and data handling resources in 1970-71. Therefore, air pollution control programs have been limited to whatever historic data had been assembled during that period.

(2) Regional and State Offices Have Used Ambient Monitoring Information and Have Acquired Stationary Source Emissions Data to Develop State Implementation Plans.

In 1971, as part of the SIP development process, regional and state offices began to engage in three activities, as follows:

- Analysis of Ambient Air Quality
- Identification of Sources of Pollution
- Development of Control Strategies

Regional and state offices evaluated the air quality of each AQCR using available data to determine pollution levels with respect to national standards. Detailed annual ambient air data were collected from existing monitoring sites to identify violations of national primary standards.

Regional offices utilized contractors and state agencies to identify sources of pollution within problem AQCR's. The objectives in building the source inventories were first, to identify all potential contributing sources, and second, to collect detailed emissions and parametric data for these sources. Because of the time constraint associated with the SIP development and the magnitude of data collection, accurate and complete information concerning the sources was often unavailable.

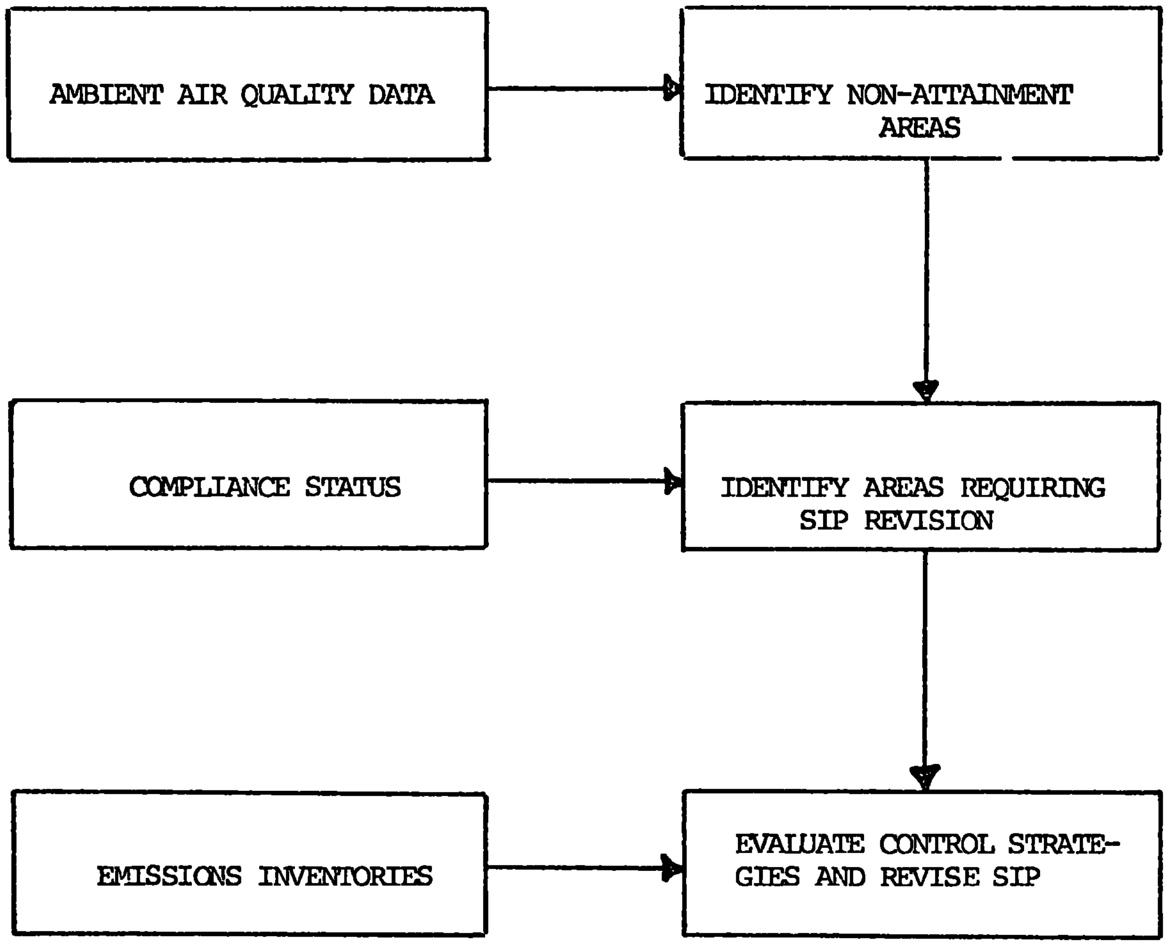
In most cases, roll-back strategies were used to develop control regulations. Under the roll-back approach an analysis was made to ascertain what percentage pollution levels should be lowered to meet national ambient air quality standards. The required reduction was applied to the emissions of the identified stationary sources and source compliance schedules were developed as a result.

(3) Regional and State Offices Need Detailed Monitoring Information and Current Emissions Data to Revise State Implementation Plans.

The FY 1976 air program guidance asked regional offices to determine the attainment status of all AQCR's for each criteria pollutant and to call for needed SIP revisions by July, 1976. Exhibit 3-3, on the following page, summarizes the data needs for this SIP revision process.

Accurate, complete and reasonably current ambient air monitoring data are needed for each criteria pollutant to identify non-attainment areas. Data must be consistent with the measurement units and the averaging time for the respective NAAQS. Non-attainment may result from the lack of an adequate SIP or from the lack of appropriate enforcement efforts under the existing plan. For this reason, accurate, current and complete source compliance information is needed also by regional offices. In addition, short-term emissions data and expanded information

Exhibit 3-3
U.S. Environmental Protection Agency
Summary of SIP Revision Process



on seasonal and diurnal patterns of source operations would be desirable to evaluate and revise SIP's. However, these short-term and diurnal data are oftentimes prohibitively expensive to acquire.

A SIP revision will culminate in the promulgation of new control regulations. Regional, state and local agencies must interact and review the restrictiveness of existing regulations before new control strategies which yield more stringent emissions limitations are developed.

With the availability of representative source and emissions data and validated models, air planners are relying less on roll-back approaches to pollution abatement. Consequently, modeling is being used increasingly to evaluate control strategies. General topographic information has a direct relationship to the effectiveness of models because of the state-of-the-art in model development. Additionally, the effective utilization of models requires detailed source parameters, emissions inventories and meteorological data to set initial conditions. Also, specific and accurate local ambient data are required to calibrate models. Modeling provides regional and state offices with the capability to project changes in air quality that may result from future source emissions or new control strategies.

In addition, regional and state offices require general research information as to the availability of cost-effective control technology. Regions need information concerning the short-term abatement capabilities of differing classes of control equipment and their associated long-term viabilities.

(4) <u>Regional and State Offices Require Localized and Summary Data to Develop Transportation Control Plans (TCP).</u>

TCP's are used primarily in metropolitan areas in order to control the aggregate emissions of hydrocarbons, nitrogen dioxide and carbon monoxide from mobile sources. Consequently, annual air quality information is needed for these three pollutants to identify urban areas where TCP's would be appropriate.

Traffic flow information, fuel availability and mobile source emissions for various model years are needed to analyze pollution problems and to develop appropriate control mechanisms. However, these data are prohibitively expensive to acquire except on a very selective basis. Localized demographic and meteorological data are needed also to support the limited modeling that is being done as a part of TCP development. For example, the State of California has projected mobile source emissions in five-year intervals from 1975 to 1995. Transportation data (CALTRANS), emissions factors and growth indices are used in conjunction with transportation models to update regional transportation plans for the urbanized areas of California. As a result of this analysis, California determined that it needed to establish stringent motor vehicle emissions standards in order to achieve its pollution reduction objectives.

2. EPA AND STATE OFFICES NEED COMPLIANCE AND SITE-SPECIFIC MONITORING INFORMATION IN ORDER TO IMPLEMENT ABATEMENT PLANS AND PROGRAMS.

Detailed and localized air quality, emissions and compliance data are needed by regional and state offices to carry out plans which will achieve the objectives of the Clean Air Act. Exhibit 3-4, on the following page, summarizes the information needs characteristics associated with the implementation of abatement programs.

(1) EPA and States Require Compliance Status and Emissions Information on Stationary Sources to Monitor the Progress in State Plan Implementation.

Exhibit 3-4
U.S. Environmental Protection Agency
Information Needs to Implement
Abatement Plans

	Ambient Air Data			Emissions Data						
	Accuracy	Currency	Completeness	Accuracy	Currency	Completeness	Compliance Data	Other Data	Summary or Detail	National or Local
Site Surveillance	N/A	N/A	N/A		Annual	10% Regions 100% State	Current Complete		D	L
SIP Enforcement	✓	✓	✓	✓	✓		Current Complete	Technical Legal	D	L
Non-Significant Deterioration Reviews		✓ Current	✓	✓	✓ Current	✓	Neighboring Sources	Meteorological Topographic Demographic	D	L
Mobile Source Emissions Performance	N/A	N/A	N/A	✓	Annual	✓	In-use Surveillance		D	N
Monitor TCP, Fuels and Vapor Recovery Programs		Annual					Inspection Results	Source Inventories	S	L
Quality Assurance Programs	✓	Annual	✓	N/A	N/A	N/A		Management Control Plans	D	L

✓ = Important characteristic
Blank represents less important characteristic
N/A = Not applicable

The successful implementation of state plans requires that regional and state offices monitor major stationary sources for compliance with established abatement schedules. Regional Surveillance and Analysis (S&A) divisions are advised by program plan guidance to inspect at least 10% of the major point sources within their jurisdiction each year. Moreover, in fulfillment of EPA grant requirements, state inspectors must visit all the major sources within their geographical area annually.

SIP monitoring requires detailed and site-specific source and emissions information along with status information on source compliance with abatement schedules. S & A divisions gather and analyze the field inspection reports and advise Enforcement personnel of suspected violations.

(2) <u>EPA and States Need Current Detailed Technical and Emissions Information to Support Enforcement Actions.</u>

Enforcement actions may result from the analysis and investigation of five sources of information, as follows:

- Compliance status reports
- Field inspection reports
- Section 114 inquiries
- Source performance reports associated with NSPS or NESHAPS
- Citizen complaints

Regional and state enforcement offices require detailed, accurate and current emissions information to conduct investigations and support case development. Contractors are used often to augment EPA efforts to collect technical data from state and plant files and to conduct comprehensive source performance tests.

Regional enforcement personnel receive technical and legal assistance from Headquarters offices, as needed. Engineering expertise concerning abatement methods, equipment and industrial processes is needed to provide technical support in enforcement actions. Legal information is obtained through oral and written communications between state, regional and Headquarters attorneys.

The Division of Stationary Source Enforcement (DSSE) at Headquarters needs various air quality and emissions information to support three activities, as follows:

- Development of national enforcement strategies
- Assessment of quarterly enforcement progress for reporting to the Administrator
- Preparation of annual reports to Congress on source compliance status

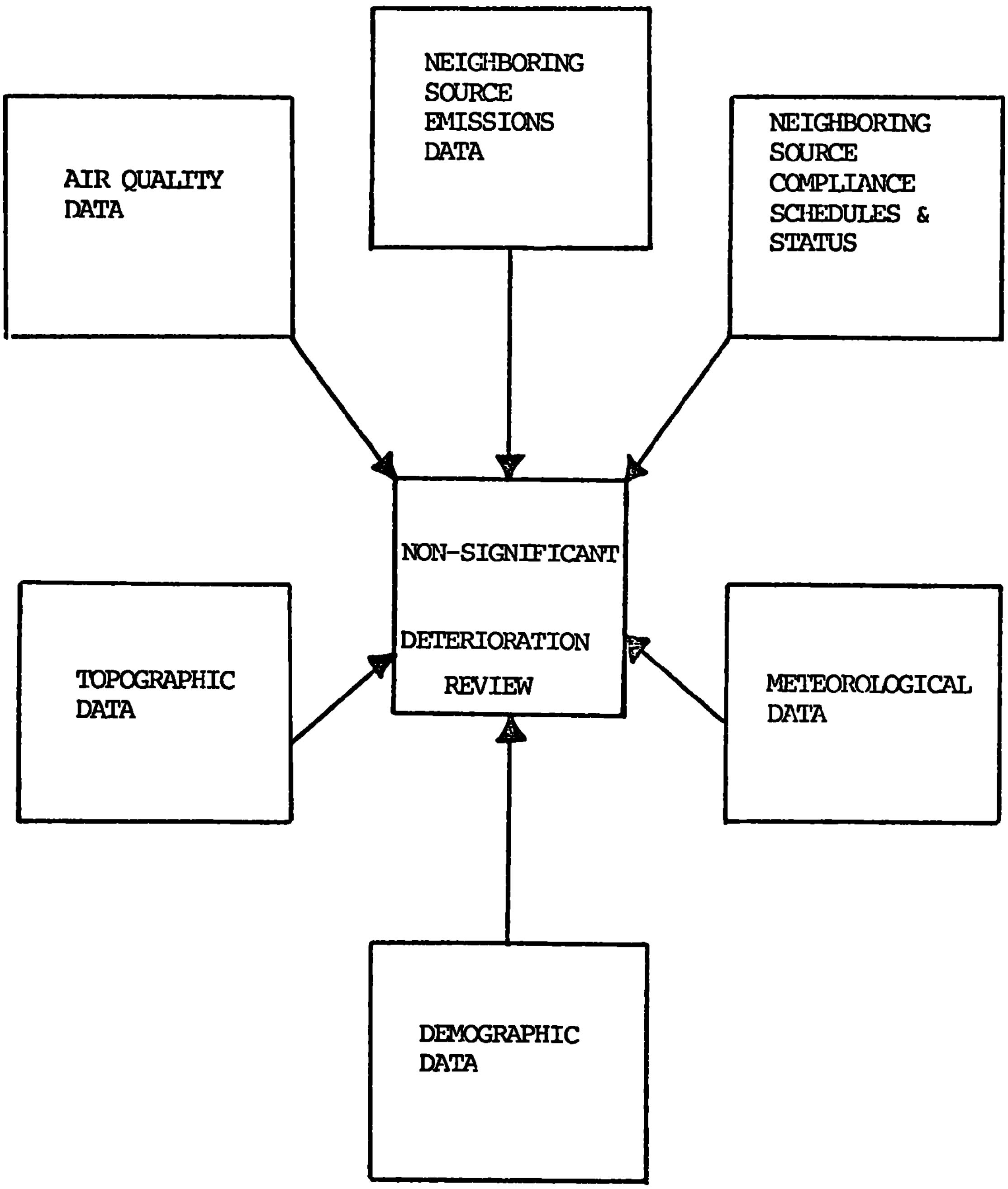
To support these efforts DSSE has noted the need for data which are no more than three-to-six months old. Furthermore, base year data are needed to assess progress in attaining goals.

(3) Regional Offices Require Ambient Monitoring, Technical Source and Emission Information to Conduct New Source Reviews (NSR).

Regional offices conduct NSR's to determine the incremental effect on air quality of a proposed stationary source. Exhibit 3-5, on the following page, summarizes the information needed for a review.

Localized and complete ambient monitoring and emission data are needed for the January, 1975 baseline and for the current period to perform analyses associated with non-significant deterioration reviews. Regional offices use models to project pollution levels and to analyze the associated increments in air quality.

Exhibit 3-5
U.S. Environmental Protection Agency
Information Needs for New Source Review



Consequently, current, complete and accurate demographic and meteorological data concerning the area surrounding the proposed source are needed as input to various models. Additionally, regional offices need compliance status and emissions history information about neighboring sources to augment and interpret model results.

(4) Prototype Test Results and In-use Monitoring Information Are Needed to Assure Mobile Source Emissions Performance.

Each year the mobile source program reviews applications from manufacturers to certify that new vehicles comply with emission standards prior to their introduction and sale. Detailed and accurate emission data are collected on 600 prototype vehicles every 5,000 miles for 50,000 miles. The test procedures generate data which are representative of the actual use of the vehicles under various normal driving conditions. EPA also needs actual use emissions data on vehicles which have been purchased and driven by owners. The results of this testing are used to support four functions as follows:

- Analysis of vehicle emissions for compliance with EPA requirements
- Calculation of emissions rates for TCP and SIP development
- Analysis of control equipment for possible enforcement orders to manufacturers for recall
- Investigation of possible emission control tampering

As a by-product of the national emission testing, EPA collects and publishes fuel economy data in cooperation with the Federal Energy Administration.

(5) Compliance Status Information Is Required to Monitor the Effective Implementation of Transportation Control Plans, Vapor Recovery and Fuels Programs.

Regional and state offices need to review annual ambient air data to ascertain the effectiveness of TCP's in reducing mobile source-related emissions of hydrocarbons, nitrogen dioxide and carbon monoxide in selected urban areas.

Complete source inventories of service stations, fuel distributors and bulk terminals are needed to schedule inspections in order to review compliance with vapor recovery regulations. When a plan is not on schedule, enforcement action can be taken to re-establish conformity with schedules.

Regional offices visit service stations to inspect for compliance with regulations concerning signs, equipment and lead content. Data are gathered from quick screening tests to find gross violations of the unleaded standard and from an atomic absorption measurement scheme for more precise quality control. Either test may be used as a basis for an enforcement action. Data from both tests are reported to Headquarters and stored in the Fuels Data Base.

(6) Management Control Plans and Analyses of Detailed Ambient Data Are Required to Support Regional and State Quality Assurance Programs for Monitoring Sites.

Air quality data produced by federal, state and local monitoring sites must be accurate and legally defensible. Accordingly, states develop quality assurance programs in conjunction with EPA personnel in order to correct seven deficiencies associated with ambient monitoring, as follows:

- Improper location of monitoring sites
- Improper placement of instruments
- Inadequate calibration of instruments
- Improper measurement methods and materials
- Inadequately trained site operators
- Cumbersome data handling
- Inadequate laboratory support facilities

Regional offices use statistical validation of air quality data and audit results of equipment calibration, data handling and operational procedures to evaluate monitoring site performance.

To support these efforts, the Monitoring and Technical Support Program within ORD, together with OAQPS, develop agency-wide quality assurance guidelines and technical information for use by regional and state offices. This guidance includes four areas, as follows:

- Standard measurement methods
- Standard analytical reference materials
- Monitoring site selection and validation criteria
- Training and technical assistance

EPA anticipates that efforts to upgrade state and local monitoring programs will result in improved data quality.

3. <u>EPA NEEDS DETAILED MONITORING, SOURCE AND METEOROLOGICAL DATA AND CLINICAL TEST RESULTS TO ANALYZE PROBLEM AREAS AND TO DEVELOP NEW PROGRAMS.</u>

Various types of engineering and scientific information are needed by Headquarters planners and research personnel to deal with specific pollution problems and to develop new control programs. Exhibit 3-6, on the following page, summarizes the information needs associated with these areas.

(1) <u>EPA Needs Detailed Non-Criteria Ambient Air and Clinical Test Data to Identify New Pollutants to Be Controlled.</u>

The Health and Ecological Effects program within ORD collects data and provides information for the establishment and subsequent reevaluation of air quality criteria. In support of this activity, detailed and accurate short-term ambient air data are needed to determine exposure-response relationships of various pollutants. Specific health effects data are collected from three sources, as follows:

- Clinical studies
- Toxicological studies
- Epidemiological studies

Exhibit 3-6
U.S. Environmental Protection Agency
Information Needs to Analyze Problem
Areas and to Develop New Programs

	Ambient Air Data			Emissions Data						
	Accuracy	Currency	Completeness	Accuracy	Currency	Completeness	Compliance Data	Other Data	Summary or Detail	National or Local
Identify New Pollutants	✓						N/A	Health Studies NASN Analysis, Demographic	D	L
Revise NAAQS	✓						N/A	Health Studies	D	L
Develop Simulation Models	✓		✓	✓		✓	N/A	Topographic Meteorological	D	L
New Source Performance Standards					Annual	✓	N/A	Control Technology	S	N
ESECA		Annual	✓		Annual	✓		Source Parameters, Fuels Availability	D	N

✓ = Important characteristic

N/A = Not applicable

blank represents less important characteristic

Most of the air quality monitoring in recent years has been for the determination of ambient concentrations of total particulate matter, sulfur dioxide, carbon monoxide and oxidants. The section of these four pollutants has been based primarily upon the state of scientific knowledge about air pollution effects.

Recently, there has been an increasing need for air quality information on non-criteria pollutants such as sulfates. Exhibit 3-7, on the following page, summarizes some important non-criteria pollutants and the associated areas of concern. Data on ambient concentrations of non-criteria pollution are collected primarily through the National Air Surveillance Network (NASN) and analyzed by the Environmental Monitoring and Support Laboratory within ORD. This federal network consists of 250 monitoring sites located around the country, primarily in urban areas.

Exhibit 3-7
U.S. Environmental Protection Agency
Summary of Non-Criteria Pollutant
Sources and Areas of Concern

POLLUTANTS	MAJOR SOURCES	GENERAL AREA OF CONCERN
Sulfates	Atmosphere reactions sulfur oxide precursors (in moist environment appears as sulfuric acid mist or rain)	Aggravation of respiratory diseases; corrosion of metals and other substances
Nitrates	Atmospheric reactions of nitrogen oxide precursors (in moist environment appears as nitric acid mist or rain)	Aggravation of respiratory and cardiovascular illnesses
Asbestos	Asbestos mining and processing	Fibrosis, calcification and cancer of the lungs
Lead	Automotive emission; lead smelting and processing	Impairment of nervous system, bones and kidneys
Mercury	Mining and refining of mercury	Impairment of nervous system
Beryllium	Mining, smelting and machining of beryllium	Suspected cause of bone and lung cancer

Non-criteria pollution data for a larger number of areas are needed to allow more complete analyses of ambient and population dosage levels.

Ambient monitoring data for pollution assessment and research must be collected under rigorous quality control conditions to assure validity and accuracy. Consequently, prior to each study, researchers need information concerning at least five areas, as follows:

- Purpose for which data are collected and used
- Number and placement of monitoring sites for spatial resolution
- Frequency of sampling
- Quality control associated with sampling and analysis
- Specific data constraints

Data needs are seldom identical from study to study; therefore, raw ambient data collected through state programs are not sufficient for the specific research needs.

(2) EPA Needs Research Information to Reevaluate Air Quality Criteria and Revise NAAQS.

On a continuing basis EPA conducts research on both the short- and long-term exposure effects on air pollution on animals, materials and vegetation. This effort provides the basis by which EPA reevaluates air quality criteria and recommends revised NAAQS. EPA's research is augmented by studies conducted by universities, other government agencies and other countries.

New research is being conducted also in various areas to reassess the adequacy of current ambient air standards. These research activities include analyses of the effects of varying sizes of particulate material, of the synergistic effects of combinations of chemicals and of the feasibility of limiting or controlling the presence of hazardous combinations of chemicals. Researchers require a wide variety of specialized statistical and data-handling tools to assist them in these efforts. In most cases the data and tools required to support these analyses are very unique and cannot be provided economically by national data bases of trend and monitoring information.

(3) <u>EPA Requires Detailed Ambient, Emissions, Meteorological and Topographic Information to Develop, Evaluate and Validate Air Quality Simulation Models.</u>

The Ecological Processes and Effects program of ORD develops analytical techniques that relate source emissions to ambient air pollution levels. The monitoring sites associated with the Regional Air Pollution Study (RAPS) in St. Louis collect detailed and accurate air quality information that can be used to calibrate and validate simulation models for criteria pollutants.

Accurate and complete emissions inventories, meteorological data and other scientific data are needed by ORD model developers and contractors. Emission and process data are required for a variety of industrial processes to develop and update emissions factors. Additionally, ambient data around large point sources, monitoring site information that allows for spatial resolution and short-term emissions data are needed to calibrate models and to validate results. Appropriate meteorological data are difficult to obtain. Information concerning wind speed and direction above and around stacks is important for diffusion modeling; however, weather information collected at high altitudes and at ground level is usually the only data available. Furthermore, topographic area characteristics are needed to provide guidelines for physical model utilization. St. Louis represents a relatively flat terrain; consequently, models developed for that area would not be appropriate for use in evaluating strategies in the more mountainous western states, for example.

(4) <u>The Office of Air Quality Planning and Standards (OAQPS) Needs Summary Stationary Source Emissions Data to Establish New Source Performance Standards.</u>

EPA has initiated an assessment program to determine which industries should be controlled by new source performance standards. This effort is complicated because of the large number of sources that can be addressed. Accordingly, EPA ranks industrial source categories by analyzing five types of information, as follows:

- National annual emissions by industrial class for specific pollutants
- Relative toxicity of the pollutants
- Population in proximity to sources
- Control technology availability, effectiveness and costs
- Physical configuration and industrial processes associated with sources

OAQPS personnel gather this data with the assistance of the Industrial Processes Program of ORD to establish emissions standards which can achieve desired levels of control economically.

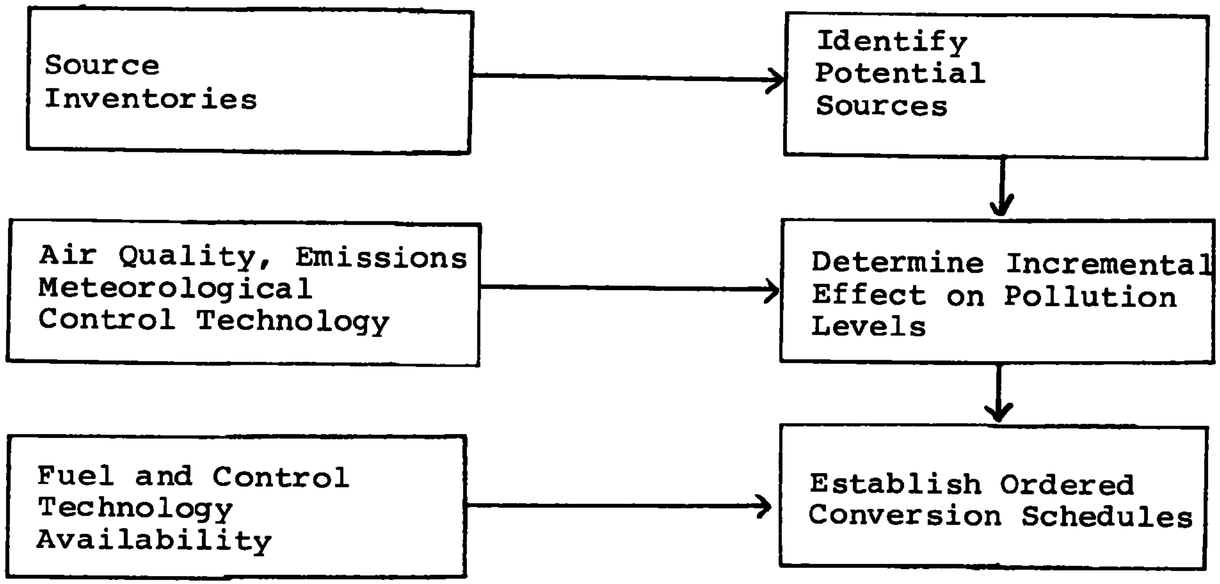
For example, in January, 1976, EPA promulgated a fluoride standard for new sources associated with the primary aluminum industry. As part of its analysis EPA considered the more stringent standards that had been established previously by the State of Oregon. There were several reasons why EPA elected not to adopt standards equivalent to those of Oregon. In particular, EPA believed that the Oregon standards would require the installation of relatively inefficient, costly secondary scrubbing systems at virtually all new primary aluminum plants. This action would have had a substantial adverse economic impact on the aluminum industry. By contrast, EPA's standard requires use of secondary control systems only for a specific classification of aluminum plants. Therefore, EPA concluded that cost considerations precluded adopting the Oregon standards as NSPS for the primary aluminum industry.

(5) <u>EPA Requires Stationary Source Inventories, Fuels Availability, and Ambient Monitoring Information to Evaluate Conversion Plans Associated With the Energy Supply and Environmental Coordination Act (ESECA).</u>

Under ESECA the Federal Energy Administration (FEA) may issue orders prohibiting power plants and other facilities from burning scarce petroleum products or natural gas. Exhibit 3-8, on the following page, summarizes the information required by EPA to analyze and plan its actions associated with ESECA.

EPA and FEA have used source inventories to identify power plants for potential conversion orders. Detailed and complete air quality, emissions and meteorological data are needed to model areas to determine the incremental effect on pollution levels that would result from conversion of these fuel-burning sources to coal. In addition, information relating to the

Exhibit 3-8
U.S. Environmental Protection Agency
Information Needs for ESECA Analysis



availability and associated transportation costs of low-sulfur coal are needed to assess the feasibility of conversion. If a conversion order is deemed feasible, then EPA and FEA use the model results, associated control technology information and demographic data to determine whether a suspension of SIP requirements is required and justifiable in order to convert the designated facility to coal.

* * * * * * * * *

In summary, air program personnel require a variety of ambient, emissions, demographic, engineering and research information. National summary ambient and emissions data are needed by EPA program planners and policy makers. Regional, state and local personnel, by contrast, require significantly more detailed local information to support SIP revision, AQMA evaluations and modeling. Regions, in particular, require detailed local source data for NSD, NSPS, NSR, NESHAPS and enforcement activities. The increasing use of models for SIP revision, new source review and transportation control strategy development requires complete localized source inventories and comprehensive meteorological, topographical and demographic information.

The availability of required data is dependent upon two factors, as follows:

- The economic and scientific feasibility of acquiring the information
- The technical data processing considerations associated with storing and reporting acquired data

In most cases the data processing costs for storing and reporting data are small compared to the costs for measuring accurately the ambient, emissions, meterological and other desired data. Program managers have made the cost determinations feasible to capture. The next chapter discusses the data processing systems which have been employed to edit, store and report on these captured data.

IV. CURRENT AIR ADP SYSTEMS

IV. CURRENT AIR ADP SYSTEMS

This chapter describes the ADP systems that support the air program. These systems fall into six areas, as follows:

- Storage and Retrieval of Aeromatic Data (SAROAD)
- National Emissions Data System (NEDS)
- Compliance Data System (CDS)
- Comprehensive Data Handling System (CDHS)
- Other Aeros Systems
- Research Systems

Each of the systems are discussed with regard to six parameters, as follows:

- History, purpose and uses
- Data flows
- Reporting capabilities
- Technical efficiencies
- System management
- System costs

More detailed technical data on each system appear in Appendix A.

1. <u>THE STORAGE AND RETRIEVAL OF AEROMATIC DATA (SAROAD) SYSTEM STORES AND REPORTS AIR QUALITY AND SAMPLING SITE DATA COLLECTED FROM MONITORING STATIONS AROUND THE COUNTRY</u>.

SAROAD operates on the UNIVAC 1110 computer at the National Computer Center (NCC) at Research Triangle Park (RTP), North Carolina. It is maintained by the National Air Data Branch (NADB) within the Office of Air Quality Planning and Standards (OAQPS). SAROAD is designed to store and report air quality-related data collected from air monitoring sites around the country. Its data base consists of monitoring site descriptive data, air quality sampling values, and summary statistics such as means, maximum values, and running averages.

(1) <u>SAROAD Is Used to Monitor, Analyze, and Plan Programs for the Abatement of Air Pollution Levels</u>.

SAROAD was developed initially in 1966 to store air quality data collected from 250 federally maintained monitoring sites associated with the National Air Surveillance Network (NASN). In 1972, NADB was created to maintain data supplied by states as part of their state implementation plans. SAROAD was selected by NADB to store these air quality data submitted by the states. As a result, the number of active sites tracked by SAROAD has increased to approximately 4,000.

Since 1966, SAROAD has undergone two major conversions to accommodate hardware changes. The latest conversion occurred in 1974-1975 when SAROAD was shifted to the UNIVAC. Problems associated with these equipment changes and hardware malfunctions have impaired greatly NADB's ability to provide SAROAD users with the level and quality of ADP service which NADB desires to provide. Currently, SAROAD is undergoing major modifications by contractors to increase throughput, to improve data storage efficiencies, and thereby to reduce the exposure of the current system to UNIVAC hardware malfunctions.

EPA uses the data stored in SAROAD for four major purposes, as follows:

- Perform trend analysis
- Determine compliance with ambient air standards
- Evaluate abatement strategies
- Support air quality maintenance planning

Users of SAROAD data have indicated that the system report formats and flexibility are generally sufficient to meet program needs. Major criticisms have been raised, however, regarding data quality, completeness and currency. These deficiencies are due primarily to the scientific complexity of the air monitoring strategy and not to the SAROAD data system itself. Regional users have noted, however, that the unstable UNIVAC hardware environment and the prolonged data edit and correction procedures associated with SAROAD-submitted data has resulted in regional and state reluctance to insure that submitted data is examined sufficiently and corrected promptly.

State usage of SAROAD is limited primarily to verifying the accuracy of data which have been submitted to SAROAD. States may request reports for analysis purposes, but usually they do not rely on the SAROAD data to perform analyses. Instead, states rely on data contained in their own local automated or manual systems.

Other government agencies and private institutions also use SAROAD information. Private industry can request information under the Freedom of Information Act. These requests are placed through the regions or through OAQPS.

(2) <u>SAROAD Data Originate Primarily From the States and Are Submitted Through the Regional Offices</u>.

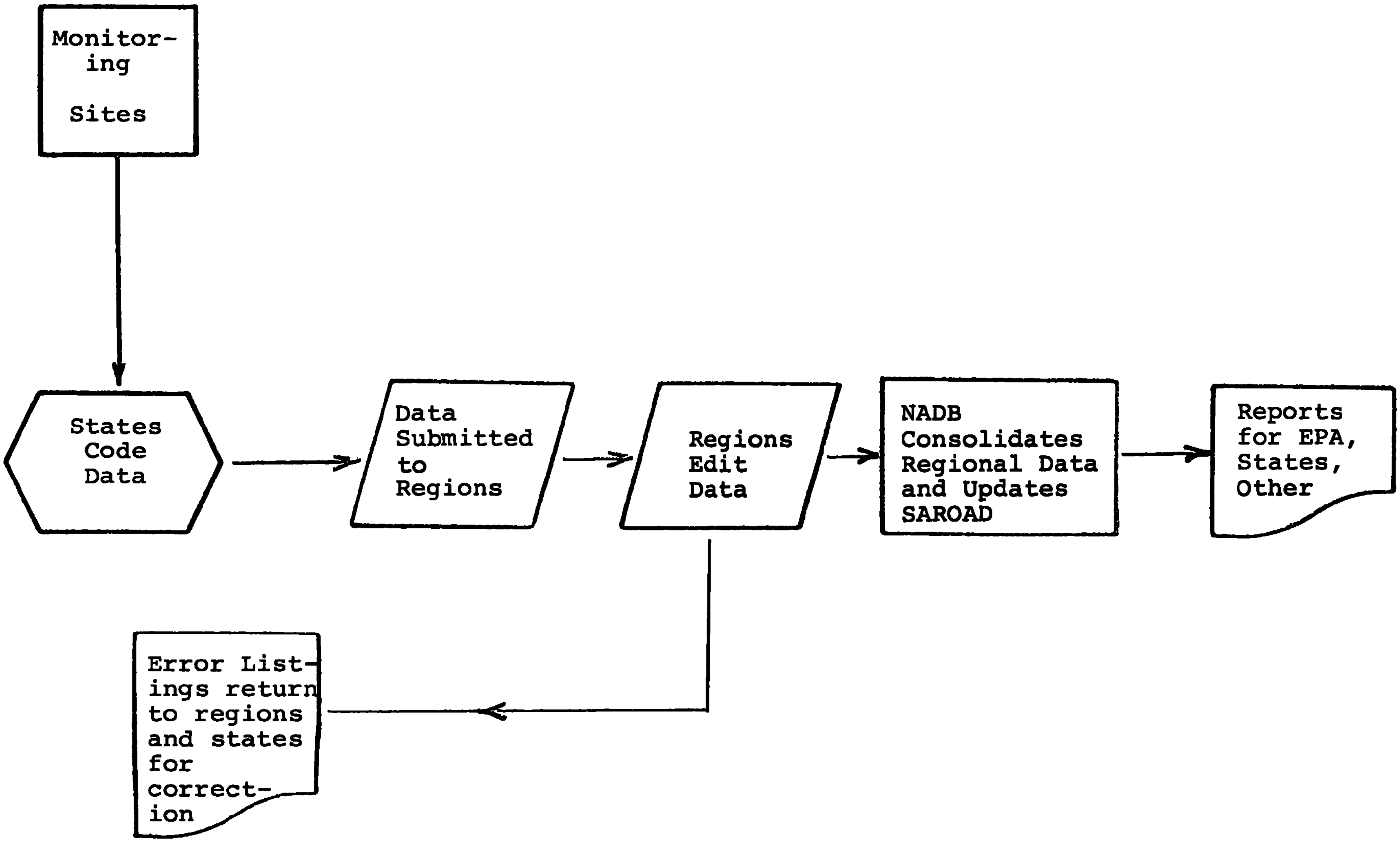
Most SAROAD data originate with the states, who are required to submit quarterly air quality data to EPA as part of their state implementation plans (SIPS). Regional offices have been designated by EPA to act as intermediaries for data transferred between the states and the SAROAD system. The regions receive state data, initiate data input and editing, and review data anomalies and errors with the states. Regional and state personnel have indicated that they are often late in submitting and correcting SAROAD data because they have insufficient personnel resources to provide all data and to correct erroneous data in a timely fashion.

States do not have access to SAROAD directly. Instead, the states must rely on regional contacts for information regarding SAROAD contents. Exhibit 4-1, on the following page, depicts the SAROAD data flow from states through the regional offices into SAROAD.

Currently, SAROAD data must pass through six steps, as follows:

- State personnel code monitoring site data
- State personnel submit data to EPA regional office
- Regional personnel run edit programs on state-submitted data
- Regional personnel correct errors or return to states for resolution
- NADB consolidate regional data and update SAROAD
- NADB generate reports for EPA and others

Exhibit 4-1
U.S. Environmental Protection Agency
SAROAD Data Flows



The states are required to submit data on sites, pollutants, and sampling values. They may submit these data on coding forms, cards, or magnetic tape, depending upon each state's data processing capabilities. Regions, in turn, input these data into a series of SAROAD edit routines which check for proper formats, valid codes, and reasonable data values. Edited data from the regions are updated periodically into the SAROAD data base. These updates are initiated only by NADB. Updates are scheduled based upon the number of data values received by NADB and upon software and hardware problems with SAROAD and the UNIVAC.

Regions usually enter data and retrieve reports via RJE terminal. In cases where regions do not have tape drives or where large quantities of data must be transmitted, the regions mail data to the National Computer Center.

(3) <u>SAROAD Produces a Wide Range of Statistical and Raw Data Reports</u>.

SAROAD can produce about 20 standard reports and also can support user-generated report programs. These user reports can be produced by employing a wide variety of selection parameters, such as geography, pollutant and year. EPA users can access reports directly via RJE or through on-line low-speed terminals. States, private industry and other federal agencies can request data through either the regional offices or OAQPS.

There are three major types of reports produced by SAROAD, as follows:

- Raw data reports
- Summary reports
- Management reports

Raw data reports can show four different types of data, as follows:

- Sampling values
- Site descriptions and sampling methodologies
- Means, maximum values, standard deviations and other arithmetic values for a specified time period
- Values which exceed standards

Summary reports do not show the actual interval values but summarize these values by quarter or year. Other descriptive information on sites and sampling methods is similar to that on the raw data reports. Summary reports include such values as minima, maxima, percentiles and standard deviations.

Management reports record the activity status of sampling stations. These reports show four types of data, as follows:

- Measurements required by EPA
- Measurements proposed in SIPS
- Measurements actually reported
- Minimum and maximum values for each quarter by site

These reports can be accessed by batch terminal access to the UNIVAC. Interactive reports can be retrieved for certain select summary reports.

(4) Headquarters and Regional Users Are Satisfied with SAROAD Reporting Capabilities.

Overall, users at OAQPS make extensive use of SAROAD reporting features. Regional users have made more limited use of SAROAD reports in the past because of problems with the UNIVAC equipment and questions regarding the quality and currency of SAROAD data. Regions have indicated, however, that SAROAD report formats meet their needs and that they will increase their report request rate as the UNIVAC hardware stabilizes. Some regions have begun to use SAROAD reports to perform more comprehensive edit checks on state-submitted data from CDHS. These regions have indicated that the superior reporting capabilities of SAROAD enable them to flag anomalous values which CDHS does not catch.

(5) <u>SAROAD Requires Substantial Computing Resources to Perform Updates and Generate Reports</u>.

SAROAD is a tape-oriented system operating on the UNIVAC 1110 at the National Computer Center. All raw data values are maintained on 26 tapes. These tapes contain about sixty million values and are expanding at the rate of twenty million values annually. For each day's worth of data there is a key containing ten parameters such as state, area, site, agency and project. These keys are used by report programs and system update routines to access specific data, as required.

In addition to raw data tapes, SAROAD also contains quarterly and yearly summary and frequency data stored on sequential disk files. These files are accessed by most summary report programs. Random access quarterly and yearly summary and frequency files are maintained also to facilitate direct access by system users. These files are accessed by the AEROS Timesharing System (ATS) to produce on-line summary reports.

A SAROAD update involves three steps, as follows:

- Updating the raw data tapes from edited input transactions

- Creating the sequential quarterly and yearly frequency and summary files
- Creating the direct access quarterly and yearly summary and frequency files

The profile of resource hours required to perform an update is shown in TABLE 1. These time estimates assume an update volume of approximately 2,500,000 data values.

TABLE 1

Step	SUP Time	
Raw data files updated	7	hours
Summary	11.5	hours
Frequency files	12.5	hours
Total	31	hours

Thus SAROAD consumes approximately 31 SUP hours of UNIVAC resources to complete an average update.

Report retrievals from the sequential tape files or sequential disk files are time consuming, since many records must be searched before a desired record is located. Total SAROAD update and report retrieval time accounts for about 27% of the annual UNIVAC resource utilization.

(5) SAROAD Currently Is Undergoing Technical Modifications to Improve Data Storage and Throughput Capabilities.

Modifications to SAROAD are underway currently to reduce data storage requirements and to improve processing efficiencies. Some modifications have been completed and others are scheduled to be finished by the beginning of fiscal year 1977.

Preliminary estimates show that these changes will reduce costs for updates and report retrievals. Seven major changes are being implemented to the system, as follows:

- Store less than 24-hour data on disk rather than tape and eliminate a large number of tape files
- Place summary and frequency files onto a random access file and eliminate current sequential files
- Create summary and frequency files directly from the raw data files rather than from intermediary tape files

- Produce running averages during system updates and store these averages within the raw data file
- Compress data from full-word to half-word, thereby reducing the data base size by 50%
- Reduce daily keys with one key for a year's worth of data, thereby saving disk space
- Produce summary and management reports from direct access random files, thereby reducing report retrieval times and eliminating tape contention problems

These changes are expected to reduce SAROAD update times from 31 SUP hours to approximately 8 SUP hours and report retrievals by a smaller percentage.

(7) <u>The National Air Data Branch Is Responsible for Maintaining SAROAD</u>.

The National Air Data Branch (NADB) in Durham, North Carolina, is responsible for maintaining SAROAD and the other AEROS systems. NADB is composed of three sections, as follows:

- Systems Development Section (SDS)
- Data Processing Section (DPS)
- Requests and Information Section (RIS)

SDS has primary responsibility for system development work, and system and user documentation. Most contractor development activities related to SAROAD and other AEROS systems are managed by this section.

DPS is responsible for collecting and processing all data submitted to EPA under state reporting requirements and for formulating regional procedures for data entry and retrievals. Under certain circumstances DPS will edit data for regional offices. In addition, DPS initiates all SAROAD updates.

RIS coordinates with SAROAD users through three activities, as follows:

- Receives, coordinates and tracks the status of all requests received by NADB, including those by regions and those under the Freedom of Information Act
- Communicates with SAROAD user contacts within each region
- Distributes AEROS newsletters and coordinates user training sessions

(8) SAROAD ADP Costs Are $903,000 for Fiscal Year 1976.

SAROAD is the most expensive computer system running on the UNIVAC. Fiscal 1976 costs, including regional and OAQPS charges, are estimated at $903,000. Exhibit 4-2, on the following page, presents a breakdown of these ADP costs.

Approximately 23% of these costs are attributable to modifications to the SAROAD system associated with changing the system design to run more efficiently within the UNIVAC environment. These costs are expected to decrease substantially in 1977. Maintenance costs of $314,000 include OAQPS editing and report generation. Report costs of $390,000 include both regional and OAQPS reporting. Contractor costs for 1976 are approximately $220,000. EPA personnel costs are difficult to estimate because the responsibilities of personnel in each of the three NADB sections include other AEROS systems also. However, the EPA manpower in support of all AEROS systems is approximately 20-25.

2. THE NATIONAL EMISSIONS DATA SYSTEM (NEDS) STORES AND RETRIEVES POINT AND AREA SOURCE EMISSIONS DATA.

The National Emissions Data System is maintained by the National Air Data Branch and is operated on the UNIVAC

Exhibit 4-2
U.S. Environmental Protection Agency
SAROAD ADP Costs

	1975	1976
Development	$ 11,000	$ 199,000*
Maintenance	263,700	314,000
Reports	331,542	390,000
TOTAL	$ 606,242	$ 903,000

* Most of these development costs are for the redesign of SAROAD to run more efficiently on the UNIVAC.

1110 at NCC. NEDS contains emissions and descriptive information on point and area sources throughout the United States. NEDS was designed and developed initially by EPA personnel to provide a basic source inventory for program planners. It currently supports trend analysis and modelling as well. Point source data were collected initially by EPA and by outside contractors. Since 1973, states have been responsible for submitting point source data to EPA semi-annually on selected sites as part of the state implementation plans. EPA continued to maintain the NEDS area source and emission factor information.

(1) NEDS Is Used to Support Modeling and Control Strategy Evaluation.

NEDS stores descriptive data, emission factors and emission estimates on 160,000 point sources and 3,200 area sources. These data are used mostly by OAQPS and regional offices.

OAQPS uses NEDS primarily for generating national emissions summaries and for refining emission factors. OAQPS employs several customized analysis programs to process NEDS data in support of these functions. Regional offices use NEDS for three additional purposes, as follows:

- Evaluating state air pollution control strategies
- Inputting, editing and reviewing state data
- Modeling

States do not use NEDS directly but are responsible for verifying their input data. Instead, states maintain local automated or manual systems for tracking local emissions data. Some NEDS data are used by other government agencies and by private institutions.

Overall, NEDS has been used primarily by OAQPS. Regions have tended to do less modeling than they would like and have used state- or source-provided data when doing modeling. Regions indicated a reluctance to use NEDS because of the incomplete and non-current data contained in the system. NADB noted that most of this incompleteness was attributable, however, to state offices which have failed to submit their data and to regional offices which have failed to enforce these reporting requirements.

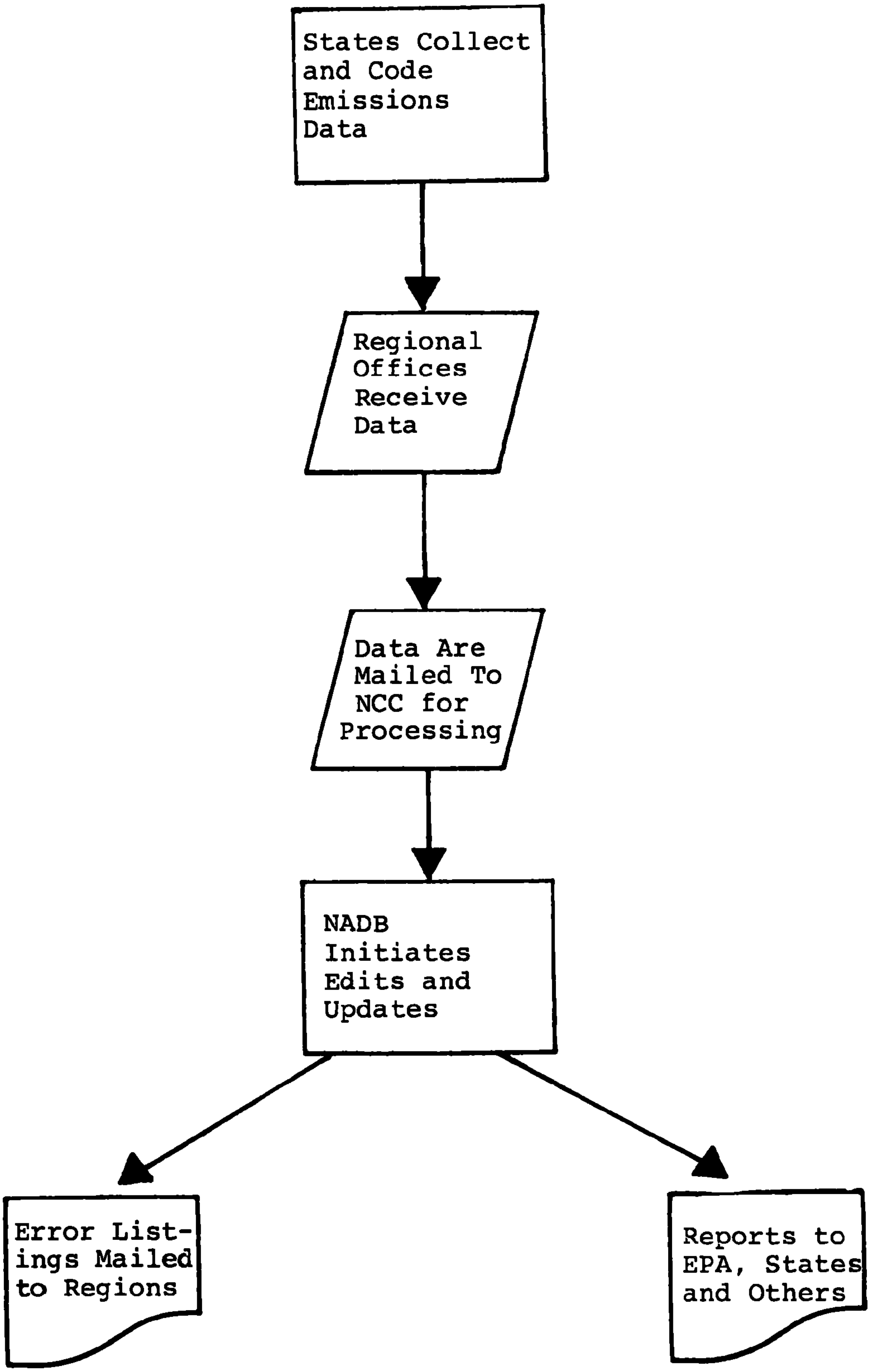
(2) <u>NEDS Data Originate Primarily With the States and Are Submitted Through the Regional Offices</u>.

Most NEDS data come from state and local agencies. States are required to maintain accurate data on all point sources which emit more than 100 tons per year of criteria pollutants.

States collect data on point sources through site visits, audits and reviews of industry-submitted documentation. Depending upon the data processing capabilities of the submitting states, these data are sent by states to regions on either coding forms, cards or tape.

After the regions receive the state data, they mail or transmit the data to NCC. NADB then edits the data and mails edit listings back to the regions. These listings are returned by the regions to the states for verification. The flow of data from states to regions to NADB and back to the regions for editing may take as long as two months. States and regions have indicated that these delays in NEDS data flows hamper their ability to assure that accurate and complete data is submitted. Because of these delays, NADB has undertaken the development of a NEDS edit for use directly by regional offices. Exhibit 4-3, on the following page, depicts the data flow for NEDS. Data edits and updates occur concurrently for NEDS. Updates are scheduled bi-weekly and are initiated only by NADB The update process consists of an update to a master tape file and the re-creation of a direct access user file.

Exhibit 4-3
U.S. Environmental Protection Agency
NEDS Data Flows



(3) <u>The NEDS Data Base Supports Point and Area Source Data Retrievals and Various Modeling and Statistical Analysis Applications</u>.

NEDS produces approximately twenty batch and on-line reports for different selection criteria. Reports can be printed locally or centrally at NCC and mailed to the users. Specialized reports also can be written by system users.

NEDS produces three types of reports, as follows:

- Raw data reports
- Summary reports
- Management reports

These reports cover both area source and point source files. Raw data reports for point sources can include five types of information, as follows:

- Plant identification and location
- Plant operating characteristics
- Stack and control equipment data
- Emission estimates
- Fuel or raw material processing rates

Area source raw data reports can contain three types of information, as follows:

- Area identification
- Emissions and activity levels of the area source processes
- Total emissions of the five criteria pollutants

Summary reports may be generated for geographic areas or by SCC codes for emission and fuel usage data. System management reports describe the data base contents, including missing values and recent updated sources.

(4) <u>NEDS Is Currently Undergoing Modifications to Improve Updating and Data Storage Efficiencies</u>.

NEDS operates on the UNIVAC 1110 at the National Computer Center. Its primary file is an indexed sequential user file occupying 1-1/2 disk packs. This file contains point and area source descriptive data and emissions estimates for each source. This file is updated from a tape containing the descriptive data and from various smaller files containing geographic parameters and SCC emission factors. NEDS calculates emissions based on three factors, as follows:

- Emission methods and estimates stored within the descriptive file
- Operating characteristics of the source
- Relevant SCC emission factors

All emissions are recalculated during each update, regardless of whether the values have changed. The current NEDS update procedure takes about two SUP hours. Modifications are underway by contractors to alter the NEDS updating procedure so that only records that change are updated in the user file. This change is expected to reduce NEDS updating costs by 50%.

(5) The National Air Data Branch Is Responsible for Maintaining NEDS.

Procedures for operating and maintaining the NEDS are similar to those for SAROAD. NADB is responsible for updating the system and for managing new system development activities. The three sections within NADB are responsible for system development, operations, user requests and engineering and scientific support. Documentation for NEDS is maintained by a central documentation group responsible for all AEROS systems. NEDS user information appears in the monthly AEROS newsletter.

(6) NEDS Computer Costs for 1976 are $150,000.

The NEDS system is expected to cost $150,000 in fiscal year 1976. Exhibit 4-4, on the following page, shows the associated costs for OAQPS and regional offices. The regional costs of $11,000 for 1976 reflect the limited regional reporting activity during most of the fiscal year. Overall costs for 1977 are not expected

Exhibit 4-4
U.S. Environmental Protection Agency
NEDS ADP Costs

	1975	1976
Development	$ 4,687	$ 32,000
Maintenance	49,573	14,000
Reporting	205,871	104,000
TOTAL	$ 260,131	$ 150,000

to increase. However, there will be a shift of some costs from OAQPS to regions as a result of the decentralization of the NEDS edits and as a result of increased regional reporting.

3. THE COMPLIANCE DATA SYSTEM (CDS) RECORDS AND REPORTS STATIONARY AIR COMPLIANCE AND ENFORCEMENT DATA.

The Compliance Data System (CDS) operates on the IBM/370 at the Optimum Systems, Inc. (OSI) facility in Maryland. The system was designed and developed in 1972-1973 by the Division of Stationary Source Enforcement (DSSE) within the Office of Enforcement. CDS provides automated support to regional and Headquarters Air Enforcement personnel. It records and retrieves information regarding source increments of progress and enforcement actions relating to SIPS, NSPS, and NESHAPS schedules. All data input, file update, and data retrievals use batch processing.

(1) CDS Supports Regional Compliance Monitoring and DSSE National Trends Analysis.

The primary users of CDS are regional office compliance clerks and compliance engineers. Program managers, and section, branch, and division chiefs also use the printouts and summary reports. In addition, DSSE personnel use system printouts prepared on a national basis for tracking enforcement progress for major industrial categories and the nation as a whole.

CDS was developed in response to the long-range stationary source air pollution program plan which prescribed a means of managing the vast amounts of facility-related compliance information to be generated by the air program. Headquarters planners defined the essential requirements to be (i) to forecast when increments of progress of compliance schedules were to be completed by any source, and (ii) to track other enforcement actions and compliance conditions. The system was not considered to be a replacement for the hardcopy files which are maintained in the regional offices but rather to be a tool for managing and summarizing the data contained in those files.

(2) Compliance Data Are Obtained From States and Are Entered Into CDS by Regional Offices.

Regional offices are responsible for gathering compliance information from state offices. Information is obtained from state and local air pollution personnel via standard reporting mechanisms, routine daily conversations, and region/state agreements on data exchange. In addition, regional engineers, attorneys, and compliance clerks gather and provide information to the regional CDS clerks for encoding. Generally, information is coded onto regionally-customized forms prior to keypunching.

The EPA regional offices are responsible for the collection, preparation and submission of data for input. The regions review, edit and update reports and correct any data errors or omissions. In addition, regions formulate, code and execute retrievals at the regional data centers for use by local personnel.

Presently, data are submitted in card image form for editing. Transactions passing the edit are collected on a storage file to be held and subsequently processed during the next update.

Once entered and verified, data remain relatively unchanged with the exception of the action-related records, compliance condition items, and any comments records. All regions are expected to keep these items current within three months.

(3) <u>CDS Produces Reports for Tracking Stationary Compliance Status and for Scheduling Enforcement Activities</u>.

The system was designed for use in tracking the status of stationary air compliance with schedules and emission regulations and for scheduling and tracking the occurrence of enforcement-related events. By providing the users with listings of events that either are to occur, have occurred, or should have occurred, the completion of an event can be noted in the system files. Enforcement action can begin and be noted on the system for those sources which are delinquent in complying with emission abatement schedules.

Four basic report formats are available from the system, as follows:

- Source Data Report
- Quick Look Report
- Action Questionnaire
- Action Summary Report

These reports are available through a customized CDS retrieval package.

(4) CDS Employs a Tape Master File and a Disk-based Retrieval Capability.

CDS was developed in COBOL for operation on an IBM/370 computer. It contains a sequential tape master file which is copied to disk after each update for use by the retrieval program. System software, source programs, load modules, procedures, files and related data sets are maintained on two IBM 3330 disk packs.

CDS provides a list of facilities, emission points, and compliance schedule actions that were to occur in the previous six-month period. Facility and emission points can be matched to NEDS records through the use of the state, county, NEDS source cross reference, and NEDS point cross reference codes.

The CDS files record data on over 30,000 major and minor air facilities throughout the country. Each week an average of 10,000 transactions are processed in the update run.

(5) CDS Is Maintained and Operated By the Division of Stationary Source Enforcement (DSSE).

DSSE is responsible for the overall operation and maintenance of CDS. DSSE secures funds and manages contractor support for system operation, program maintenance, system enhancement, data coding and keypunching assistance. In addition, DSSE acts as a central focal point for solving problems with system operation and for coordinating and approving suggestions for improvement. Modifications and enhancements to the system are managed by DSSE under the recommendation of an annual Headquarters/Regional user meeting.

(6) <u>CDS Computer Costs for Fiscal Year 1976 Are $105,000</u>.

CDS computer costs are divided into regional and DSSE processing. The table below shows manpower requirements and ADP contractor costs for 1975 and 1976 for both regions and Headquarters.

	Computer	Personnel	Contractors
<u>Headquarters</u>			
FY 75	$ 50,000	1	$ 75,000
FY 76	$ 54,000	1	$ 75,000
<u>Regions</u>			
FY 75	$ 30,000	5	$ 125,000
FY 76	$ 51,000	5	$ 125,000

Regional contractor costs are for data encoding. These costs are expected to decline slightly for fiscal year 1977.

4. THE COMPREHENSIVE DATA HANDLING SYSTEM (CDHS) SUPPORTS STATE EMISSION INVENTORY AND AIR QUALITY DATA PROCESSING REQUIREMENTS.

CDHS is a combination of two separate computer systems, the Air Quality Data Handling System (AQDHS-II) and the Emission Inventory System (EIS/PR). AQDHS-II handles air quality monitoring data and EIS processes emission inventory, permit and registration data. AQDHS-II is designed to interface with SAROAD, and EIS/PR is designed to interface with NEDS. Both systems were developed by EPA with outside contractor assistance and are implemented on state ADP facilities with EPA contract funds. Currently, 25 AQDHS-II and 21 EIS/PR systems are operational or are scheduled for installation. Either system can be operated independently of the other.

(1) CDHS Supports State Collection, Editing and Retrieval of Air Quality and Emissions Data.

In 1973, EIS/PR and AQDHS-II were developed by EPA to support the data handling and reporting requirements of state agencies and to perform SAROAD and NEDS functions, respectively, at the state level. EIS/PR also contains permit and registration information and local site data that are not contained in NEDS. These systems support three functions, as follows:

- Edit and store aerometric data
- Provide reporting capabilities for retrieving information from the data base
- Produce card or tape formatted output that can be used as input into NEDS and SAROAD

As of March, 1976, eleven states were using AQDHS-II and fourteen states were using EIS. EPA managers estimate that the number of states using these systems will increase by 100%-150% over the next two years. Currently, five-to-ten states maintain their own local systems for performing AQDHS-II and EIS/PR functions. The remaining states use manual filing procedures.

Occasionally, regions request data from states to compare state data against those found in EPA systems or to correct currency or completeness deficiencies in the EPA data bases.

(2) <u>AQDHS and EIS Receive Data From State Air Pollution Agencies and Produce Output Data for EPA AEROS Systems</u>.

Data gathering methods of EIS/PR and AQDHS-II data are similar to those of NEDS and SAROAD. States collect emission inventory data from site audits and from data supplied by point sources and submit those data into EIS. Air quality data from selected monitoring sites operated within the states are collected, encoded, keypunched and entered into AQDHS-II. Data can be loaded also from SAROAD or NEDS into the AQDHS-II or EIS/PR data basis, respectively. For EIS/PR, area source data can be taken from NEDS, or can be input manually by the states.

Both AQDHS-II and EIS maintain edit routines to check for proper formatting and to identify data anomalies. Data which have passed edit routines successfully are updated into the master files. These updates are initiated by the individual states or local agencies responsible for maintaining the systems.

EIS/PR and AQDHS-II output information on cards or tape for input to SAROAD and NEDS. These outputs are in proper format for processing by the EPA systems. However, since the AQDHS-II and EIS/PR edits have not been identical to those of SAROAD and NEDS, not all CDHS transactions have been accepted by the EPA systems. SAROAD and NEDS edit listings are returned to the regions and subsequently are returned to the state for verification. Since comparisons between the state and

EPA systems must be done in a time-consuming manual fashion, states often do not reconcile all data inconsistencies between the local systems and NEDS and SAROAD. For example, EIS/PR provides no automated method, at present, of monitoring which data have been submitted to NEDS and which transactions still must be posted to reconcile the two systems.
By contrast, AQDHS-II maintains internal flags which indicate the specific transactions which have been used to generate SAROAD update transactions. However, some regional offices submit all AQDHS-II and EIS/PR data into the EPA systems periodically, rather than attempt to identify the specific change transactions which would be needed to reconcile the differences between the data bases.

(3) <u>AQDHS-II and EIS/PR Offer Standard Reporting and Specialized Programming Capabilities</u>.

AQDHS-II and EIS/PR produce standard reports but users can also produce customized reports by using COBOL or the CDHS report-generator package. There are four standard AQDHS-II reports, as follows:

- Detailed list reports by sampling interval
- Sliding average reports, which include the number of readings, mean and maximum values

- Data analysis reports, which display arithmetic values and percentiles for various selection criteria
- Formatted SAROAD output file dumps

There are three standard reports for EIS/PR, as follows:

- File lists, which show one point source or area source record per page
- Summary reports on emissions data across various selection criteria
- Formatted NEDS output file dumps

These standard reports are enhanced by user-generated report retrieval programs which provide additional information needed by state managers.

(4) <u>CDHS Systems Are Designed to Operate on Different Computer Configurations While Maintaining Standard Edit, Report and File Maintenance Procedures</u>.

AQDHS-II and EIS/PR are designed to operate on both tape and disk files. The systems are written and have been installed on a variety of different computers including IBM, UNIVAC and CDC equipment.

EIS/PR is composed of three major components, as follows:

- Master file
- Edit, update and retrieval routines
- Emission factors table

The master file can reside on either tape or disk. It contains files similar to those of NEDS, although the files are accessed only by plant segment, point source, or fuel and process keys. The plant segment contains descriptive data on the site, including codes for geographical and plant identification. Each point source contains information on operating characteristics, compliance status, control equipment and emissions. The fuel and process keys contain data on process output and emission factors. The EIS/PR master file also contains area source parameters similar to those used within NEDS.

The EIS master file is constructed from information supplied on add, change or delete transactions. Unlike NEDS, which only tracks emissions data on five pollutants, EIS/PR is capable of storing data on as many as sixteen different pollutants. This enables states to maintain emissions inventories on non-criteria pollutants, such as lead or flouride, which may be included in state regulations.

EIS/PR also contains a special report-generator program which allows the user to program special reports. In addition, EIS/PR has routines for converting EIS/PR point and area source data into NEDS input format and for converting NEDS data into EIS/PR readable formats.

AQDHS-II contains many of the same features as does EIS/PR. AQDHS-II maintains a similar report retrieval capability, and can convert data to SAROAD input format. The AQDHS-II data base also contains data similar to those which are maintained by SAROAD on monitoring stations, sampling intervals and sampling values. However, there are no summary or frequency files maintained in AQDHS-II. Furthermore, CDHS does not provide for on-line retrieval of information.

Editing capabilities for both EIS/PR and AQDHS-II are more limited than those available from NEDS and SAROAD. CDHS edits are adequate for screening out most formatting errors but do not provide good tools for identifying unrealistic input values.

(5) AQDHS-II and EIS/PR Are Maintained and Installed by EPA and Contractors But Are Operated by State Agencies.

The CDHS systems were first conceived in 1970 and were developed with EPA funds during the period 1972-1974. The first systems were installed in late 1973. Systems generally are installed by EPA contractors on computer systems used by the states. Installation costs average $17,000 and include a one-week user training session. Once a system is installed, the state is responsible for subsequent operating costs associated with the system.

The CDHS system software is maintained by NADB. An EPA contractor is available to perform trouble-shooting, when necessary, and to provide additional system enhancements for implementation by the states. Currently, there is one person responsible for coordinating enhancements, programming problems, and user interfacing. Periodically, NADB distributes a CDHS newsletter to state and regional users which informs users of changes in system capabilities.

The contract costs for AQDHS-II and EIS/PR are difficult to identify separately since some contracts involve work related to both systems. Estimated contractor costs for CDHS system development and maintenance, user support and documentation are shown in Exhibit 4-5, on the following page. They amount to $521,000 for fiscal year 1976. Seventy-two percent of these costs are related to implementing the systems within the states. The remainder is associated with development and continuing user support. NADB personnel responsible for CDHS estimate that $400,000 will be needed over the next two years to continue system development and user support. Another $100-200,000 are expected to be spent on state implementations. This estimate is based on an additional five-to-ten systems being implemented in the states over the next two years.

Actual computer processing costs for CDHS are difficult to determine precisely, since usage varies with state size and reporting needs, and state costs differ depending upon state hardware and billing algorithms. However, CDHS personnel estimate that 12-24 CPU hours are used annually by small states, and 24-48 CPU hours for large states, at an average of $300 per CPU hour.

Exhibit 4-5
U.S. Environmental Protection Agency
CDHS Contractor Costs

	1975	1976	1977-1978 (2 year estimate)
State Implementations	$123,000	$380,000	$100-200,000
Development & Enhancements	$109,000	$ 96,000	$ 250,000
User Support & Documentation	$ 55,000	$ 20,000	$ 50,000
Trouble-Shooting	$ 50,000	$ 25,000	$ 100,000
TOTAL	$337,000	$521,000	$500-600,000

(6) <u>The Enforcement Management System (EMS) Is Being Replaced by State Systems Compatible with the Compliance Data System (CDS)</u>.

EMS was developed in 1972 for use by state and local agencies. Its purpose was to monitor the compliance status of point sources. It is installed, at present, in five states. The EMS system design was used as the basis for the Compliance Data System, which is operated by the Office of Enforcement. However, EMS currently cannot interface easily with CDS. Instead, the Division of Stationary Source Enforcement (DSSE) is adapting CDS for use by states to facilitate the interface between state data and the EPA version of CDS. There are no plans to expand the capabilities of EMS or to install it in additional locations.

5. <u>OTHER AEROS SYSTEMS ARE USED TO RETRIEVE AIR QUALITY OR EMISSION DATA, AND TO PERFORM STATISTICAL ANALYSES</u>.

There are ten other AEROS systems which are used to support air program activities, as follows:

- Energy Data System
- Form 67
- Quality Assurance Management Information System (QAMIS)

- State Implementation Planning System (SIPS)
- Source Test Data System (SOTDAT)
- Hazardous and Trace Elements Inventory System (HATREMS)
- Weighted Sensitivity Analysis Program (WSAP)
- Regional Emission Projection System (REPS)
- Source Inventory and Emission Factor Analysis (SIEFA)
- Computer-Assisted Area Source Emissions Guiding Procedure (CAASE)

These systems are designed to operate on the UNIVAC 1110, and in many cases, access data from either SAROAD or NEDS.

(1) The Energy Data System (EDS) and the Federal Power Commission Form 67 System Report Energy-Related Data on Power Plants.

EDS and Form 67 are two separate computer systems used by EPA for enforcement, modeling and planning. EDS is being developed currently with contractor support to operate on the UNIVAC 1110. Form 67 is maintained by the Federal Power Commission. EPA receives a copy of the data base annually and operates the system on the UNIVAC.

The Strategies and Air Standards Division (SASD) has developed EDS to aid in assessing the air quality and energy impacts of environmental legislation. EDS provides interactive access to energy-related data collected from other EPA data banks. EDS receives its data from NEDS, SAROAD, CDS and Federal Power Commission Form 67. In addition, emission regulations and environmental policy data are entered manually into the system. When fully developed, the data base will contain seven types of data, as follows:

- Fuel quantity and consumption data
- Plant design and operating data
- Emission regulations
- Compliance information
- Future megawatt capacities
- Diffusion modeling results
- Air quality data

EDS produces interactive reports that cover five major areas, as follows:

- Fuel use summaries by geographical region and by categories
- Emission and equipment information for large fuel-burning sources
- Compliance schedules and status

- Modeling results for large power plants
- Air quality data in the vicinity of large power plants

ADP costs for EDS for fiscal year 1976 are approximately $186,000. EDS zero-based budget estimates for 1977 are $300,000. EDS requires two man-years of EPA staff to maintain the system and data base. Development contractor obligations since 1974 have amounted to $402,000. These contract monies include work scheduled for completion through 1977.

Since 1969, the Federal Power Commission (FPC) has collected information on monthly and annual fuel use by boiler, and associated sulfur and ash content from each power plant over 25 megawatt capacity. This information is stored in the FPC's Form 67 system. EPA receives magnetic tapes from FPC each year for inclusion in its data base. Approximately 400 items are submitted annually for each power plant. Currently, there are about 800 power plants in the file. Form 67 can report power plant information according to various selection criteria and sort options. Fiscal year 1976 computer costs are about $175,000. Contractor costs for 1976 and 1977 are expected to be $160,000.

(2) The Quality Assurance Management Information System (QAMIS) Stores and Retrieves Quality Control Data on Air Pollution Collection Agencies.

QAMIS was designed in 1974 by the National Air Data Branch to record questionnaire data submitted to agencies, and laboratories involved in monitoring, collecting, and interpreting air quality data. Data were collected on approximately 250 agencies and laboratories which collectively operated and maintained 2,500 monitoring sites. The purpose of the data was to evaluate the existing air quality monitoring networks which were submitting data to SAROAD. QAMIS provides the ability to analyze quality assurance information relating to SAROAD sites, agencies and measurement methods.

QAMIS produces reports by five standard selection criteria, as follows:

- Agency
- Laboratory
- Site
- Pollutant
- Opinion survey

The QAMIS system has not been updated since 1974 and there are no current plans by either OAQPS or ORD to expand the system. Current annual costs for computer usage are approximately $2,000.

(3) The State Implementation Planning System (SIPS) Stores and Reports Text on State Implementation Plans.

In 1974, the SIPS system was designed by NADB as a tool for storing, retrieving and analyzing text associated with state implementation plans. With the assistance of the Division of Stationary Source Enforcement and regional offices, NADB input approximately 5,000 state regulations into the SIPS data base. Parameter codes were established for legal authorities, control strategies, geographic areas and other identifying criteria. Currently, NADB and the Standards Implementation Branch within OAQPS are verifying and updating the text. The primary source of current changes and modifications to SIPS are taken from the Federal Register.

SIPS produces listings upon request of selected portions of state implementation plans. Approximately 18 selection codes can be used to access specific regulatory information. In addition, reports can be produced according to various selection criteria, such as state, region, regulation status and date. SIPS contractor development and maintenance costs for 1976 are $38,000 and are expected to increase slightly over the next two years. Manpower support for SIPS is currently one-to-one-and-a-half manyears and is expected to stabalize at two manyears in the future. Computer costs for fiscal year 1976 will be approximately $13,000.

(4) <u>The Source Test Data (SOTDAT) System Stores, Analyzes and Retrieves Stationary Source Test Data</u>.

SOTDAT operates on the UNIVAC 1110 at the National Computer Center and is maintained and operated by the National Air Data Branch. This system was developed beginning in 1973, with contractor assistance, and became operational in January, 1976. SOTDAT stores and retrieves data gathered from stack tests and other emission measurements. SOTDAT data differ from those stored in NEDS in that SOTDAT data represent actual stack emissions data, whereas NEDS data reflect estimated emission information only. Currently, there are about 700 point source test records stored within SOTDAT. These records contain four major types of information stored on each test source, as follows:

- Plant and fuel data
- Process team parameters
- Pollutant test results
- Control device data

The primary purpose of SOTDAT is to provide a means of updating and improving emission factors for specific SCC codes. In addition, SOTDAT also provides data for evaluating equipment performance and for providing construction site cost estimates. Furthermore, SOTDAT can be used to store information related to enforcement activities.

There are three major types of report retrieval capabilities within the SOTDAT system. Reports can be produced by three criteria, as follows:

- Data for all test runs by pollutant
- Emission factors by SCC code
- Descriptive data on source sites that had associated test runs

These reports can be retrieved on-line or in batch mode.

Source tests are conducted by both EPA and non-EPA sources for both enforcement and research purposes. Stack test data are collected and coded by OAQPS staff or contractors. Data sources include both published articles as well as test results submitted directly to OAQPS from other EPA offices or outside sources. The current data base is expected to increase by approximately 400-to-500 point source records annually. It is estimated that two manyears of effort are required to gather and encode data and to update the SOTDAT system. Computer costs for fiscal year 1976 are estimated at approximately $38,000. This usage level is expected to remain stable over the next several years.

(5) The Hazardous and Trace Elements Inventory System (HATREMS) Is Designed to Store and Retrieve Operating Parameters and Estimated Emissions Data for Non-Criteria Pollutants.

HATREMS is being developed currently by OAQPS and contractors to store and report emissions information on non-criteria pollutants. HATREMS is expected to be operational during fiscal year 1977. It is designed to store point source descriptive data, emission factors, and SCC data for non-criteria pollutants. The data base will have a similar structure to that of NEDS in that it will maintain both a point source and area source file. The majority of source sites in HATREMS will be taken directly from NEDS. It is estimated that an additional 100-to-500 point sources will have to be input to make a complete inventory of non-criteria pollutant sources.

HATREMS will be used by OAQPS staff to monitor and project non-criteria pollutant emissions and to develop emission factors for those pollutants. The initial point source data will be taken directly from NEDS. It is estimated that this system will cost between $25,000 and $50,000 annually and require approximately one manyear to maintain. The system will operate on the UNIVAC 1110 and will be accessible by batch processing only. Anticipated contractor costs for 1976 are approximately $50,000 and are expected to remain constant for 1977.

(6) Four Analysis Packages Are Used by EPA to Perform Emission Forecasts and Statistical Analyses Using NEDS or SAROAD-Related Data.

NADB maintains four small reporting and analysis packages that perform emission projections and analyze existing emission data, as follows:

- Weighted Sensitivity Analysis Program (WSAP)
- Regional Emission Projection System (REPS)
- Source Inventory and Emission Factor Analysis (SIEFA)
- Computer-Assisted Area Source Emissions Guiding Procedure (CAASE)

WSAP allows users to analyze quantitatively the effects of potential errors in air pollution source emissions inventories. The primary objective of WSAP is to calculate the maximum variation in emissions for each of the component sub-categories of the emission inventory. An acceptable error in total emissions of a criteria pollutant is specified for a geographical area. WSAP then calculates the maximum variance in emissions for each sub-category in order to produce the acceptable error rate for the total inventory.

REPS is a computerized air pollution emissions projection model. It combines national and regional economic forecasts with NEDS point and area source emission inventories for AQCR's. It projects annual air pollution levels for the five criteria pollutants up to the year 2000. REPS produces projections based on three parameters, as follows:

- Calculation of growth factors from economic and demographic forecasts

- Projection of emissions from base year inventories
- Inclusion of the effects of emission control devices on emission levels

REPS uses three primary sources for economic and demographic forecasts, as follows:

- EPA developed national economic growth projections from the Strategic Environmental Assessment System (SEAS)
- The Department of Commerce regional activity projections (OBERS)
- The NEDS data base

SIEFA is an analysis package that estimates the imprecisions associated with the NEDS inventory for any given geographic area. Internal SIEFA tables store pre-calculated errors and variances associated with NEDS emission factors. NEDS area and point source files are used to estimate the precision of the associated emissions. A major part of the SIEFA system is devoted to estimating the precision associated with mobile sources. The SIEFA report includes the calculated precision and accuracy of emissions for each source category. Precision is reported as the standard deviation of the emission estimate. The accuracy is estimated by counting those point sources within a source category which had inadequate data for NEDS to calculate emissions.

SIEFA substitutes standard values for these missing data items and estimates the resultant source emission values. These values are added to the NEDS emissions to present an approximate indication of the accuracy of the NEDS inventory. SIEFA reports can be printed by specifying a criteria pollutant and a geographical area.

CAASE is used to restructure the NEDS county-organized area source data into uniform grid areas. These grid formats are necessary input to three mathematical models, as follows:

- Implementation Planning Program (IPP)
- Comprehensive Dispersion Model (CDM)
- Air Quality Dispersion Model (AQDM)

CAASE apportions fuel and emissions data from the NEDS area file based on demographic characteristics within the respective areas. It provides activity levels and emissions for each criteria pollutant.

REPS, SIEFA, CAASE, and WSAP are routines which access the NEDS data base. Overall NEDS computer costs include costs for these four programs.

6. DATA SYSTEMS ARE USED WITHIN THE OFFICE OF RESEARCH AND DEVELOPMENT TO SUPPORT HEALTH EFFECTS RESEARCH, TO DEVELOP AIR QUALITY MODELS AND TO ANALYZE NON-CRITERIA POLLUTANTS.

The Office of Research and Development conducts various studies which require large quantities of highly-reliable data. Most of the ADP support related to air pollution research is associated with three programs, as follows:

- Community Health Environmental Surveillance System (CHESS)/Community Health Air Monitoring Program (CHAMP)
- Regional Air Pollution Study (RAPS)
- National Air Surveillance Network (NASN)

The principal data banks associated with each system are maintained at the National Computing Center in RTP. ORD use of these systems represents approximately 35% of current UNIVAC 1110 utilization.

(1) The Community Health Air Monitoring Program (CHAMP) and the Community Health Environmental Surveillance System (CHESS) Provide Information on the Effects of Specific Air Pollutants on Human Health.

The CHESS program was initiated in 1970 to gather air quality and health information from 22 communities located in five metropolitan areas around the country. Air quality data were collected over a five-year period for various pollutants including total and respiratory particulates, sulfur dioxide and nitrogen oxides. In addition, health information was collected from individuals and families residing within two miles of the air monitoring stations. Epidemiologic studies associated with CHESS included three areas, as follows:

- Chronic respiratory disease
- Acute lower respiratory disease
- Asthmatic attacks

The CHESS system consists of 240 COBOL and FORTRAN programs which were used to store and analyze the daily averaged air quality data and health questionnaire responses. Four types of reports are available from the system, as follows:

- Air quality summaries
- Demographic information
- Acute, episode and respiratory disease summaries
- Statistical analyses

Currently, EPA researchers are analyzing the data and proceeding with the publication of results of the various CHESS studies.

In 1973, the CHAMP program was initiated to gather data on short-term human exposures to various air pollutants. The system collects air quality from 23 air monitoring stations located in five communities. Presently, the CHAMP system stores air quality date for eight pollutants, as follows:

- Nitric oxide
- Nitrogen dioxide
- Ozone
- Sulfur dioxide
- Methane and non-methane hydrocarbons
- Carbon monoxide
- Peroxyacetylnitrate (PAN)
- Suspended particulate matter (total and respirable)

In addition, each monitoring site collects meteorological information on wind speed and direction, temperature, humidity and barometric pressure.

The principal distinction between CHESS and CHAMP is the sophistication of the monitoring network and data flow. CHAMP has been designed to acquire continuous air quality information through the use of mini-computer-controlled monitoring stations and to transmit the data to the National Computing Center (NCC) via telephone lines. In addition, each field mini-computer monitors various system parameters against present limits to facilitate equipment calibration.

Fiscal year 1976 computer charges for CHAMP/CHESS are $342,000. These costs are expected to remain stable for fiscal year 1977.

(2) The Regional Air Pollution Study (RAPS) Stores Emissions and Air Quality Data for Use by Researchers in the Development, Evaluation and Validation of Criteria Pollutant Air Quality Models.

The RAPS program was initiated in FY 1973 to collect detailed air quality and emissions data on the criteria pollutants to be used by researchers in the development and validation of urban dispersion/diffusion models. The St. Louis AQCR was selected for study under the RAPS program because of its relatively flat terrain and its isolation from other major sources of pollution.

The Meteorology Laboratory at RTP within the Office of Research and Development has overall responsibility for the RAPS project with specific responsibility for collecting ambient air quality data and for operating the Regional Air Monitoring System (RAMS). Twenty-five mini-computer-controlled monitoring stations located around St. Louis comprise the RAMS network. Sampling is done every half-second and voltages from each station are communicated to a central mini-computer in St. Louis where a daily tape of one-minute averages is created. A copy of this tape is sent to RTP for data validation and archiving of hourly averages.

In addition, RAPS requires a detailed and comprehensive source and emission inventory. The National Air Data Branch (NADB) has developed the emission inventory techniques and is participating in the collection of source/emission data.

NCC charges associated with RAPS for FY 76 are expected to be $250,000. $38,000 of these costs are attributable to OAQPS emissions inventory data collection. The remainder is attributable to ORD. ORD costs are expected to remain constant for fiscal year 1977. OAQPS costs will be eliminated.

(3) The NASN System Is Used by Research Personnel to Store and Analyze Air Quality Data on Non-Criteria Pollutants.

The Environmental Monitoring and Support Laboratory (EMSL) maintains and reports data on ambient levels of non-criteria pollutants collected from 250 EPA maintained NASN sites generally located in urban areas. In 1966, the SAROAD system was designed to store air quality data collected by the NASN network. With the creation of the National Air Data Branch in 1972, The SAROAD system increased in scope to include air monitoring data reported to EPA by state and local agencies.

EMSL continues to collect non-criteria data from NASN and others. Filters from the various NASN sites are sent to regional offices where they are weighed and forwarded to EMSL for laboratory analysis. Air quality data associated with NASN and other EMSL programs are stored in a data bank similar in format and design to SAROAD. This system, referred to as "mini-SAROAD", utilizes specialized edit and report programs, although its data base structure is essentially the same as SAROAD. In 1974, the Filter Bank System (FBS) was designed to assist the laboratory analysis performed by EMSL. Various filter information and test results are entered into FBS on-line. Approximately every two weeks a transaction tape of filter analysis data is created by FBS for an update of "mini-SAROAD".

Subsequently, "mini-SAROAD" is used for validation of data, report retrieval and non-criteria pollutant analysis. On a periodic basis validated data are provided to NADB for inclusion in the national SAROAD data base.

Total NCC charges for FBS and "mini-SAROAD" in FY 76 are expected to be $54,000.

*   *   *   *

In summary, air data systems operate primarily on the UNIVAC 1110 in support of program personnel within OAQPS and ORD. State personnel are supported primarily by CDHS in those states which have operational versions of CDHS, by locally developed systems or by manual record-keeping systems. Regional personnel make limited use of NEDS and SAROAD reports, and expend most of their air data processing resources in entering and correcting state-supplied ambient and emissions data.

V. PROBLEM AREAS

V. PROBLEM AREAS

This chapter presents an analysis of the problem areas associated with ADP support for the air program. The chapter is divided into five sections, as follows:

- Systems Deficiencies and Scientific Limitations Hindering Air Program Data Activities
- Data Flow Problems
- Technical and Operations Problems
- Cost-related Problems
- System Management Problems

1. THE SCIENTIFIC COMPLEXITY OF CURRENT MONITORING ACTIVITIES LIMITS THE USEFULNESS OF CAPTURED DATA WHICH ARE STORED IN THE AEROS SYSTEMS.

Completeness, accuracy and currency problems associated with air quality and emissions data arise from the complex siting and quality assurance problems associated with EPA's monitoring strategy. These monitoring strategy problems limit the usefulness of the data systems, NEDS and SAROAD in supporting modeling applications, significant deterioration review, trend analysis and compliance monitoring,since the information contained in the data systems can be no better than the data acquired as part of the monitoring strategy.

(1) <u>The Incompleteness of State-submitted Air Quality Data Limits the Scope of Trend and Compliance Analysis</u>.

EPA air program personnel need to evaluate air quality and emissions trends and compliance with standards in order to perform three planning and review functions, as follows:

- Review national abatement plan progress
- Develop and rank new national air program activities
- Prepare the national report for Congress and the general public on the condition and trends in air quality and emissions

However, limitations in the availability of SO_2 and TSP air quality data and the absence of good quality nation-wide ambient motor vehicle pollutant data have constrained EPA's ability to perform these functions.

EPA has established two criteria definitions concerning the ambient data reported from a monitoring station, as follows:

- Minimal year of data - at least three 24-hour monitor samples or 400 hourly values from continuous samples
- Valid year of data - at least five 24-hour monitor samples or 75 percent of the possible hourly values for a continous monitor in each quarter of a calendar year

Using these definitions, the Council on Environmental Quality (CEQ) noted that only 15% of TSP stations and 46% of SO_2 stations reporting minimal data submitted data sufficiently complete to allow for calculation of a valid annual mean as needed to review and analyze program progress. Approximately half of the reporting stations could not be evaluated for compliance with annual standards for these pollutants.

Furthermore, CEQ found that states had failed to submit sufficient data to determine the attainment status of each AQCR by pollutant. This determination requires that at least one station report a valid year of data in order to evaluate annual standard compliance and that at least one station report a minimal year of data in order to evaluate short-term standard compliance. In 1974, 37% of the AQCR's reported insufficient ambient SO_2 data to determine annual standard attainment and 15% of the AQCR's reported insufficient data to assess compliance with any ambient SO_2 standard.

In addition, SAROAD received only limited historical monitoring data for oxidants, carbon monoxide and nitrogen dioxide, as needed to measure progress in air pollution abatement of motor vehicle-related pollutants. California is the primary source of most of the historical air quality oxidant data. Consequently, most analyses are limited to the California area. Because of changes in measurement methods for nitrogen dioxide, no areas have sufficient historical data to assess NO_2 trends. As a result, the historical monitoring data submitted to and therefore available from SAROAD have proven inadequate to assess national trends in oxidant formation or in developing corresponding control strategies.

(2) <u>Incomplete and Inaccurate State-submitted Emissions Data Limit the Usefulness of NEDS-provided Information</u>.

National emissions analysis and trend reporting are hampered by completeness, currency and accuracy problems associated with the source information submitted by states to NEDS data base. In contrast to the direct measurement used to report ambient air quality, emissions estimates are the result of calculations obtained by applying emission factors to source inventory information. The accuracy, currency and completeness of NEDS point source information is deficient because states fail to notify Regions regarding new sources that come into operation, cease operation, modify operation or change fuels. Over 50% of the source records in NEDS have not been updated since 1972 because of the low priority given to submitting updated emissions data by states.

In the absence of good quality emissions data, analysts have projected emission levels on the basis of industry forecasts, fuel utilization and economic indicators. These different estimating techniques, data sources, working assumptions and definitions have produced conflicting emissions estimates and have exposed EPA and state agencies to legal attack from affected companies.

(3) The Siting and Selection of Monitoring and Emissions Information Does Not Satisfy the Data Needs of Modelers and Researchers.

Simulation models are being used increasingly by state and regional offices to support new source reviews, SIP revisions, and air quality maintenance planning activities. The models that are available currently for use are concerned primarily with non-reactive pollutants such as TSP, SO_2 and CO. Since most monitoring stations are sited for reasons other than collecting data for model calibration, most air quality data are only partially adequate for calibration purposes. Submitted emissions data available from NEDS are often inappropriate for use in modeling as well. As a result, modelers use whatever ambient and emissions data have been assembled in AEROS as part of the monitoring network strategy as a first pass. They then supplement these data with specially acquired data to improve the accuracy and validity of their modeling results.

Similarly, researchers are not always able to rely upon the monitoring network data stored in SAROAD to support special research projects for three reasons, as follows:

- National monitoring stations are sited primarily to provide a general index to pollution levels and not to satisfy specific research requirements.
- The quality of the collected national monitoring data is unknown in many cases.
- The time interval over which air quality data are sampled and averaged is not appropriate for various research studies.

As a result, research programs have established their own monitoring networks to supply specialized monitoring information to support their research projects. For example, the CHAMP program has selected certain municipalities in which to study the health impact of exposure to specific pollutants. The siting of stations within these municipalities is far more concentrated than would be needed to monitor SIP progress or to perform related trend analysis.

In addition, both CHAMP and RAPS utilize mini-computers to control and calibrate the various instruments used to acquire air quality data. By contrast, because of cost considerations, most state and local monitoring networks rely on human operators to record data from strip charts and to perform various calibration efforts.

2. DEFICIENCIES IN NEDS AND SAROAD DATA FLOWS CONTRIBUTE TO DATA INACCURACIES AND CAUSE UNNECESSARY MANUAL PROCESSING.

NEDS and SAROAD data originate primarily with states and then are forwarded to EPA regional offices for entry into the respective EPA systems. Numerous bottlenecks and inefficiencies occur during these data flows.

(1) NEDS Data Flow Problems Stem From Cumbersome Procedures Involving Multiple Organizations.

The flow of emissions data among states, regions and NEDS is complicated by five factors, as follows:

- Lengthy turnaround time for edits
- The natural complexity of emissions data and the NEDS input forms
- Need for technical expertise at the regional level to review state submittals
- State procedures and organization whereby the recorders of data (engineers) do not use the data but view data recording as an externally imposed burden
- Late state submittals

These factors affect the reliability of data submitted to NEDS, and contribute to increased workloads for regional and state offices. For example, regions generally do not receive edit feedback on NEDS data for three-to-four weeks after submittal. Corrections to these data then require an additional three-to-four weeks. Given the reluctance of state engineers to "waste time" submitting or correcting NEDS forms, these time delays contribute to the high incidence of late or non-submittals.

Similarly, filling out portions of the NEDS input form requires specialized engineering or technical expertise. However, state offices have only limited trained manpower available to perform these functions. As a result, state clerical staff oftentimes are used to fill out portions of the form. However, these clerical personnel have difficulty filling in the correct SCC and method codes. This problem is compounded by the fact that NEDS does not maintain adequate SCC codes for all plant types. NADB staff have estimated that 20% of all SCC codes within NEDS are assigned incorrectly to plants, either because of improper recording by state personnel, or because no valid SCC code exists for the plant.

Regional offices also require technical expertise to review state input and to identify incorrect or suspicious data values. In addition, regions must record which data have been submitted to NEDS. This record-keeping is necessary to prevent the redundant entry of NEDS data and to prevent the entry of data which are out of sequence.

NEDS data completeness and currency are affected also by late state submittals. By regulation, states are required to submit NEDS data within forty-five days after each semi-annual reporting period. In practice, states are frequently late in submitting required data. NADB personnel have estimated that as much as 65%-70% of states submit data late or submit incomplete data. Only one-third of the states meet the complete reporting requirements on time.

(2) <u>SAROAD Data Flow Problems Stem Primarily From Univac Hardware Problems and Related System Design Inefficiencies</u>.

The flow of ambient data among states, regions and Headquarters is constrained by three factors, as follows:

- Unstable hardware environment
- Resulting infrequent system updates
- Resulting slow report turnaround time

SAROAD has been updated only three times between December, 1975 and June, 1976. The infrequency of these updates has caused SAROAD data to be incomplete and out of date. Data base inaccuracies are due also to slow report turnaround times for regional offices which wish to verify their SAROAD data. These delays have been caused primarily by the instability of the UNIVAC, by contention problems with data stored on tape, and by system design characteristics which required lengthy program runs which were more susceptible to hardware failures.

3. COMPUTER HARDWARE AND SYSTEM DESIGN PROBLEMS HAVE REDUCED THE USEFULNESS OF AIR DATA SYSTEMS.

Hardware instability of the UNIVAC together with various data system design deficiencies have contributed to inefficient computer resource utilization and have limited the accessibility of stored data.

(1) The Instability of the UNIVAC System Environment Has Created Problems for All Users of Air Data Systems.

Since its installation in 1974, hardware problems at NCC have created an unstable system environment which has inhibited timely data processing operations and has limited the accessibility of the information stored in the principle air data systems. Frequent system crashes and prolonged downtime have resulted in three problems, as follows:

- Data destruction
- Increased processing backlogs
- Delayed software conversion associated with hardware transition

These problems have affected principally the ADP operations associated with ORD and OAQPS which together represent approximately 90% of the utilization of UNIVAC.

Exhibit 5-1, on the following page, depicts recent monthly NCC performance. The UNIVAC configuration performed relatively better during the period October through December, 1975. However, in January, 1976, the equipment performance deteriorated substantially. The high failure rate that has been experienced since January is primarily attributable to the incorporation of new hardware into the UNIVAC system on January 5. These system failures have been attributable to five general causes, as follows:

- Direct hardware difficulties (memory, CAU, disk subsystems and computer console malfunctions)
- Software deficiencies

Exhibit 5-1
U.S. Environmental Protection Agency
NCC Monthly Performance

Month	MTBF[1]	#Failures	Total Jobs		SUP Hours[2]	
			Batch	Demand	Batch	Demand
Oct 75	16.3	32	11094	10294	822	181
Nov 75	13.7	33	9477	9137	685	163
Dec 75	15.6	29	9781	9574	682	161
Jan 76	9.6	57	11624	13520	782	220
Feb 76	8.2	65	11420	12418	854	203
March 76	10.2	59	13648	16039	997	302
April 76	10.5	54	12001	13951	864	252
May 76	10.4	50	11603	14040	799	251

1 Mean Time Between Failures.

2 Standard Unit of Processing (SUP) includes CPU, I/O times and executive requests.

- Computer and computer-room environment-related problems or failures
- Procedural and operational stops
- Other unidentified causes

In recent months, actions have been initiated to correct hardware malfunctions. Certain hardware problems have been isolated and traced to the disk subsystems and temperature problems in the computer room. Operating system software problems have been identified also and have been communicated to UNIVAC for resolution. Although some improvements have been implemented, the mean time between failure is still below the NCC goal of 15 hours.

(2) SAROAD System Design and UNIVAC Operating Problems Have Contributed to Delays in Updating and Retrieving Air Quality Data.

Over the past year SAROAD updates and retrievals have been delayed because of three factors, as follows:

- Inadequate SAROAD program conversion from IBM to UNIVAC hardware
- Inefficient data base design
- Cumbersome data files

These factors together with the hardware instability of the UNIVAC have contributed to operations and scheduling problems at NCC. For example, because SAROAD raw data have been stored on over 18 tape reels, the processing of data file updates and retrievals has necessitated sequential scheduling of jobs to avoid data file contention problems. These scheduling requirements together with frequent system crashes have contributed to system backlogs and poor turnaround time for users.

Many of the program modifications for SAROAD associated with the IBM-to-UNIVAC hardware conversion were performed inadequately by EPA contractors. Approximately 75% of the programs that were converted from PL/1 and MARK IV to COBOL had to be modified again by EPA before the UNIVAC SAROAD could be installed.

Prior to system modification to UNIVAC hardware, SAROAD operated as a disk-based system with the data base residing on 11 disk packs. Because of the growth potential of the data base, NCC recommended that NADB transform SAROAD into a tape-oriented system. The data base has since more than doubled to over 25 tapes.

The SAROAD file structure and size have resulted in an excessive amount of computer time needed to update and retrieve raw data due to file contention problems. Furthermore, these long processing runs have made SAROAD susceptible to UNIVAC system crashes. Over the past year data file updates and summary file creation frequently have utilized over 40 SUP hours. As a result of frequent UNIVAC crashes, NADB has required an average of three-to-four weeks to complete routine data base maintenance functions.

(3) SAROAD Design and UNIVAC Operating Problems Have Forced EPA to Implement Significant System Modifications.

Because of the serious design and operating problems associated with running the poorly-converted SAROAD on the UNIVAC 1110, EPA engaged Integrated Services, Inc. (ISI) to undertake a major modification effort for SAROAD. After careful analysis of data file structure and content, ISI developed a scheme for a compacted and technically efficient data base involving five design changes, as follows:

- Separate the raw data file by interval code (less than 24-hour data and daily data)
- Archive raw data creating a historical file (1957-1973) and a current file (1974-present)
- Create a compact key structure storing one year's worth of data in half-words
- Utilize fixed length records for less than 24-hour data

. Utilize random access devices for less than 24-hour data

The ISI SAROAD enhancement project is in the final system testing phase. The preliminary results of the system test demonstrate a dramatic reduction in computer utilization for hourly raw data retrievals. For example, raw data listings for all 1975 ambient oxidant data formerly took four SUP hours, but have taken only five SUP minutes when run off the new disk version. These changes will eliminate also the need to sequence retrieval requests. As a result, typical turnaround time will be reduced from two-to-three days to one day. ISI has projected that the new data file update and summary file creation will utilize less than eight SUP hours compared to the current average of 31 SUP hours. Overall, these modifications will make SAROAD less susceptible to hardware crashes and will make the data more easily accessible for update and retrieval activities.

(4) <u>NEDS System Design and UNIVAC Operating Problems Have Contributed to Data Accuracy, Reporting and System Performance Problems</u>.

Over the past year, NEDS technical performance and data quality have been affected by three factors, as follows:

- Inadequate NEDS report program conversion from IBM to UNIVAC
- Inefficient system updating routines
- Limited editing capability

These factors have made it difficult for regions and NADB to maintain timely and accurate NEDS data. At present, regions cannot edit their NEDS data directly but must rely upon NADB to initiate data edits. Consequently, regions experience a delay of two-to-three weeks, on average, before receiving error reports. Notwithstanding this delay, the NEDS data may still be in error since the edit routines do not validate incoming data for reasonableness. These delays and limited editing capabilities contribute to data accuracy problems and require additional effort to resolve.

Many of the retrieval programs associated with the IBM-to-UNIVAC conversion were not converted adequately. Programs were changed from PL/1 to COBOL using a convertor program which did not document programs and which, in some cases, generated codes which did not run properly. In addition, NADB analysts found that subsequent enhancements could be performed more easily by re-coding the converted programs than by trying to patch the converted code. In order to rectify the problem, NADB used contractor assistance to modify and enhance NEDS programs, and to prepare user documentation to facilitate regional NEDS use. However, regions had difficulty executing report programs while these modifications were being implemented. These problems were due, in part, to NADB and regional pressure to provide regional reporting capabilities as quickly as possible.

Despite the conversion and subsequent contractor efforts, NEDS update procedures are still inefficient. After every update, the entire user file is recreated. This process would be more efficient if only those records which have changed since the last update were modified. Although this inefficiency does not affect NEDS data flows or data accuracy, it does contribute to excess processing cost since it takes as long to update one record as it does to update the entire file. Processing time reductions could be achieved by modifying the update to modify only those records which change. NADB has issued a contract to correct this problem.

(5) <u>CDHS Systems Require Minor Enhancement to Meet State and EPA Needs</u>.

The CDHS systems were designed to give states retrieval and analysis capabilities similar to those offered by NEDS and SAROAD, and to facilitate state submittals of air quality and emissions data to EPA. However, CDHS has three limitations which reduce the effectiveness of these systems, as follows:

- CDHS provides few standard reports and requires states to develop their own customized reports.
- User documentation is too technical for many state personnel.
- Only limited editing capabilities are provided by the systems.

CDHS was designed to offer states versatile reporting capabilities through the use of a report generator package. However, most state and local agencies have only limited ADP personnel available. As a result, they have not been able to develop as many local reports as they would prefer. Many of these states have indicated that they would prefer EPA to develop more standard CDHS reports to relieve state agencies from this report design burden.

Although states are supplied with a training course and system documentation, state managers have indicated that CDHS user documentation is generally too technical for state employees who are involved with coding, data entry or report generation. These managers have stated that they require more complete and descriptive documentation designed for use by less technically-oriented state personnel.

Manpower expenditures for EIS/PR and AQDHS-II data entry and correction could be reduced if CDHS edits were more comprehensive and if CDHS/AEROS interfaces were controlled more effectively. For example, EIS/PR does not flag data that have been sent to NEDS. Consequently, states track these transmittals manually or resubmit data in order to insure that NEDS data will be current. Region IV, for example, has instituted a policy of resubmitting all EIS data to NEDS rather than to attempt to track the specific data items which have been updated successfully in the past. These policies require more state and regional manpower and ADP expenditures than should be needed otherwise.

(6) Interface Incompatibilities Between CDHS and EPA Systems Contribute to Data Incompleteness and Non-Currency.

The CDHS systems (EIS/PR and AQDHS-II) are designed to facilitate the transmittal of air quality and emissions data to EPA air systems. However, technical features of CDHS affect data completeness, accuracy and currency.

Data which are entered successfully into a state system may be rejected subsequently when they are submitted to NEDS or SAROAD. This occurs because CDHS and AEROS do not employ the same edit and validation checks. As a result, state, regional and Headquarters personnel all become involved in modifying and correcting the EPA system data. Region IV, for example, has identified 24 distinct steps which must occur in order to adjust for data errors which result from these edit and validation differences. As a result, the currency and completeness of AEROS data suffers since the correction process can take many months or more to complete.

(7) The System Objectives and Design of NEDS and SAROAD Have Made Them Inappropriate for Supporting Selected New Program Requirements.

Not all special studies and programs have been supported adequately by the fixed set of reports and large, cumbersome data bases associated with the principal AEROS systems. NEDS and SAROAD were designed to provide standard reports and not to be used as a universal data base management system. Consequently, various other systems have been developed to fulfill specific needs of air program personnel.

The Energy Data System (EDS) has been developed as a tool for SASD planners to use in determining air quality and energy impacts associated with environmental legislation. Much of the data stored in EDS is extracted from existing data systems such as NEDS, SAROAD and CDS. EDS was designed to satisfy the need for easier access and retrieval of emissions, fuel consumption and air quality data than are available from the individual systems.

Other systems have been developed also to fulfill other emissions-related data needs. Since NEDS was designed to store calculated emissions data for the five criteria pollutants only, NADB is developing the Hazardous and Trace Emissions System (HATREMS) as an adjunct to NEDS in order to maintain and report non-criteria emissions data. Similarly, the Source Test Data System (SOTDAT) was developed to supply EPA engineers with more detailed information with which to update and improve emission factors than can be provided by NEDS.

4. <u>AIR ADP SYSTEMS COSTS OF OVER THREE MILLION DOLLARS IN 1976 ARE EXPECTED TO BE REDUCED IN 1977.</u>

EPA ADP system costs for the entry, storage and retrieval of aerometric data are expected to be $3,126,000 in fiscal year 1976. These costs include the operating costs for state CDHS systems. Exhibit 5-2, on the following page, shows air system costs for 1975 and 1976 for the four principal EPA users, as follows:

- The Office of Air Quality Planning and Standards (OAQPS)
- The Division of Stationary Source Enforcement (DSSE)
- The Office of Research and Development
- The Regional Offices

These costs were taken from the Time-Sharing Services Management System (TSSMS). The 1976 costs were annualized from the June TSSMS report.

State costs are difficult to ascertain directly, since state usage varies with data base size, reporting usage and billing algorithms for each state computer facility. The aggregate state CDHS costs of $300,000 are based on annual computer resource estimates provided by NADB system managers. The annual state computer utilization estimates were extended by the number of new state installations and

Exhibit 5-2
U.S. Environmental Protection Agency
ADP Air Costs by EPA User

	1976	1975
OAQPS	1,656,000	1,299,205
ORD	857,000	590,822
OE	54,000	48,000
Regions	300,000	277,080
States	300,000	Not Available
TOTAL	3,126,000	2,215,107

were based on an average computer billing rate of $300 per hour.

Exhibit 5-3, on the following page, shows a breakdown of costs by major system. SAROAD accounts for $903,000 of total air system costs for fiscal year 1976. This figure includes an estimated $55,000 of regional SAROAD usage during the year and also $199,000 in SAROAD development costs. SAROAD costs attributable to OAQPS account for 49% of the OAQPS computing costs, and total SAROAD costs account for 28% of all air systems costs. OAQPS costs are expected to decrease in 1977 by $300,000. Sixty percent of this savings will be attributable to reporting and maintenance due to current SAROAD enhancements. The remaining 40% will be due to a reduction in total system development work.

ORD costs are attributable primarily to the Regional Air Pollution Study (RAPS), the Community Health Air Monitoring Program (CHAMP) and the Community Health Environmental Serveillance System (CHESS). Combined ORD expenditures for these three systems amount to $597,000 for fiscal year 1976. Modeling efforts and analysis of data collected from the NASN network accounts for the bulk of the remaining $260,000 in ORD expenditures. Overall costs for ORD systems are expected to increase for fiscal year 1977, although exact figures could not be estimated by ORD system managers.

Exhibit 5-3
U.S. Environmental Protection Agency
Air Data System Costs

	1976	1975
SAROAD (includes regions)	903,000	606,242
NEDS	150,000	260,131
CDS	105,000	48,000
RAPS	250,000	117,600
CHESS/CHAMPS	347,000	289,756
EDS	183,000	18,000
CDHS	300,000	Not Available
NASN	54,000	89,000
Other	834,000	875,378
TOTAL	3,126,000	2,215,107

RAPS is expected to increase slightly. CHAMPS is still undergoing development work but operating costs should increase.

The Division of Stationary Source Enforcement maintains the Compliance Data System (CDS) and has spent $54,000 to operate the system during 1976. In addition, $51,000 has been spent by regional offices for data input, reporting and related air enforcement computing. These costs are expected to remain constant.

Because of various technical inefficiencies resulting from faulty IBM-to-UNIVAC conversion, air system costs have been higher than necessary. As a result, EPA has provided funding for various system modifications which will have the affect of reducing the related ADP costs.

5. DEFICIENCIES IN THE MANAGEMENT OF SYSTEM DEVELOPMENT AND OPERATIONS HAVE LIMITED THE EFFECTIVENESS OF AIR DATA SYSTEMS.

The development and utilization of air data systems by EPA and state offices have been hindered by management deficiencies in three areas, as follows:

- Ambiguous system management responsibilities
- Non-rigorous management of ADP usage at NCC
- Inadequate management of the UNIVAC computer and related conversion efforts

These deficiencies have limited the effectiveness of various air data systems.

(1) <u>The Lack of Clearly Defined System Management Responsibilities Within OAQPS Contributes to Inefficient Use of Computing Resources</u>.

The management of ADP contractors and system development efforts is a complex task which requires full-time monitoring by personnel with a sound foundation in ADP management principles. The National Air Data Branch (NADB) is the organization within OAQPS with the greatest experience in system development and management. Nonetheless, some system development within OAQPS has been managed by other OAQPS divisions. For example, the Energy Data System (EDS) was designed and developed, principally, under the direction of the Strategies and Air Standards Division (SASD) within OAQPS. Contrary to other system development work within OAQPS, such as the SIPS project, NADB undertook only a secondary review role concerning the EDS feasibility study, system design and system interfaces. The primary development responsibility rested with SASD personnel who were not as familiar with system management principles as were personnel within NADB.

The resulting data system, EDS, has proven extremely expensive to develop and contains a variety of features which make it very costly to operate. In retrospect, it appears that an adequate feasibility study was not undertaken for EDS. The EDS contractor was not required to define specific data needs nor was he required to analyze the costs and benefits associated with all relevant alternatives. Rather, a preliminary design was prepared which addressed functional capabilities, availability of source data and an implementation approach. The missing system justification and cost benefit analysis is precisely what more experienced ADP managers would have required of their contractors. As a result, the current EDS does not make efficient use of computing resources both because its functional design is no longer appropriate for serving management needs efficiently and because it utilizes costly on-line storage and retrieval of non-current redundant data. For these reasons, OAQPS would have been served better by limiting the management of new projects and contractors to those personnel within OAQPS with the most experience and training in ADP management practices and policies.

(2) ORD and OAQPS Will Need to Tighten Their Budgetary Control Over NCC Utilization.

Historically, the EPA computing center at RTP served ORD and OAQPS users exclusively and was regarded essentially as a free resource by these offices. Recently, the Management Information and Data Systems Division (MIDSD) has prepared a plan to develop the NCC into an Agency-wide resource subject to stricter budgetary controls. Consequently, all Agency ADP users will be allowed to shift their data processing to the UNIVAC as their ADP sub-allowances and NCC resources permit. Since the UNIVAC will be priced competitively to the IBM computer at OSI, MIDSD anticipates increased pressure for UNIVAC support from EPA program offices. Furthermore, Agency-wide ADP budgets have been severely constrained by Congress. Therefore, the program offices will be less free to shift program funds into their ADP sub-allowances to make up for ADP overruns. As a result, competition for UNIVAC resources will increase and not all air program requests necessarily will be satisfied. Rather, OAQPS and ORD will need to manage their ADP utilization more carefully than in the past to preserve sufficient ADP funds to support high priority programs. This will require increased management attention and technical expertise to insure that current and future applications operate as efficiently as possible.

(3) <u>Inadequate Management of the UNIVAC Computer and Related Conversion Activities Has Created Major Operating Problems For NCC Users</u>.

MIDSD is responsible for operating the UNIVAC 1110 at NCC yet it has failed to demonstrate any long-term improvement in UNIVAC operating performance. This has occurred despite the fact that other UNIVAC 1110 installations have been operated without comparable downtime. During late 1975, the UNIVAC finally began to meet established operating performance objectives. However, rather than continuing to operate the machine at this level, MIDSD decided to expand the equipment configuration despite the fact that such expansion in the past consistently had resulted in a deterioration in operating system performance. As a result, the UNIVAC has never returned to its 1975 performance level.

Major conversion efforts undertaken in transferring AEROS systems to the UNIVAC also indicate serious historic deficiencies in NCC planning and technical judgement. The line for line conversion approach adopted is an extremely dangerous method for transferring complex file-sensitive systems such as NEDS and SAROAD. Instead, significant redesign analysis should have been undertaken at conversion time to insure that the resulting systems would operate efficiently under the new UNIVAC environment.

* * * * * * * * * * * * * *

In summary, EPA's primary air data systems do not satisfy all user requirements in a timely and efficient manner. For the most part, these deficiencies are attributable to limitations in the data available from the current monitoring networks and to hardware and operating problems associated with the current UNIVAC computing environment. In addition, however, the current size of the major air computer data bases makes them more susceptible to hardware failures. The large volume of data submitted to these systems also has proven cumbersome to state and regional personnel. As a result, state and regional submitters have not expended as much effort as necessary to submit and confirm the accuracy of all collected data.

The Standing Air Monitoring Work Group (SAMWG) has undertaken a project to revise the requirements for state submittal of data. Preliminary results from SAMWG indicate that EPA will reduce substantially the volume of data which it will require from states. This change in monitoring strategy and reporting requirements will permit a change in air data processing strategy whereby EPA could maintain AEROS systems containing far fewer data values than at present. This change in data processing strategy could reduce AEROS system sensitivity to hardware problems and could free state and regional personnel to insure the quality and accuracy of the much smaller volume of submitted data.

The next chapter examines the implications of such a change in data processing strategy. It examines also the implications of other changes to current systems and practices.

VI. EVALUATION OF ALTERNATIVE ACTIONS

VI. EVALUATION OF ALTERNATIVE ACTIONS

The preceding chapters have described the air program strategy, information needs, current ADP support and associated problem areas. This chapter presents an evaluation of alternative actions to deal with the identified problem areas. This analysis is divided into four sections, as follows:

- Data system design philosophy for managing ambient air quality data
- Data system design philosophies for emissions and compliance data systems
- Technical data system actions to improve functional service and data system efficiencies
- Changes in ADP management responsibilities

Cost data related to the analyses performed in this chapter are contained in Appendix B.

1. A HYBRID COMPUTER SYSTEM DESIGN APPROACH FOR STORING AIR QUALITY DATA IN CONFORMITY WITH THE REVISED MONITORING STRATEGY CAN IMPROVE SERVICE TO EPA AND CAN REDUCE ASSOCIATED OPERATING COSTS.

Cost can be reduced and service can be improved if EPA adopts a hybrid system design philosophy for the management of air quality data in conformity with the revisions to EPA's monitoring strategy. Exhibit 6-1, on the following page, contrasts the impact of the

Exhibit 6-1
U.S. Environmental Protection Agency
Comparison Summary of System
Design Alternatives

	Utility	Hybrid	Current
Functional Service	Improved data handling with good availability of state-collected air quality data	Improved data handling with good availability of monitoring trend data	Good data availability and reporting flexi-bility but poor data handling
Annual Operating Cost Impact	$625,000	$600,000	$800,000
Development and Installation	$600,000 - $1,000,000	$380,000	$330,000
Risk	High technical and cost risk	Low technical and cost risk	Moderate technical risk
Management Considerations	System management problems with state dependence upon EPA for ADP support	Less complex system management	Less dependence of EPA and states upon each other

current design approach with a utility and a hybrid design approach. The hybrid alternative is less costly and provides better service at less risk than do the other approaches. The hybrid system approach reflects the programmatic shift in monitoring strategy as being developed by the Agency's Standing Air Monitoring Work Group (SAMWG). This programmatic shift will require far less raw ambient data for Headquarters and regional use and will free states and local agencies to allocate more monitoring resources for special-purpose local projects.

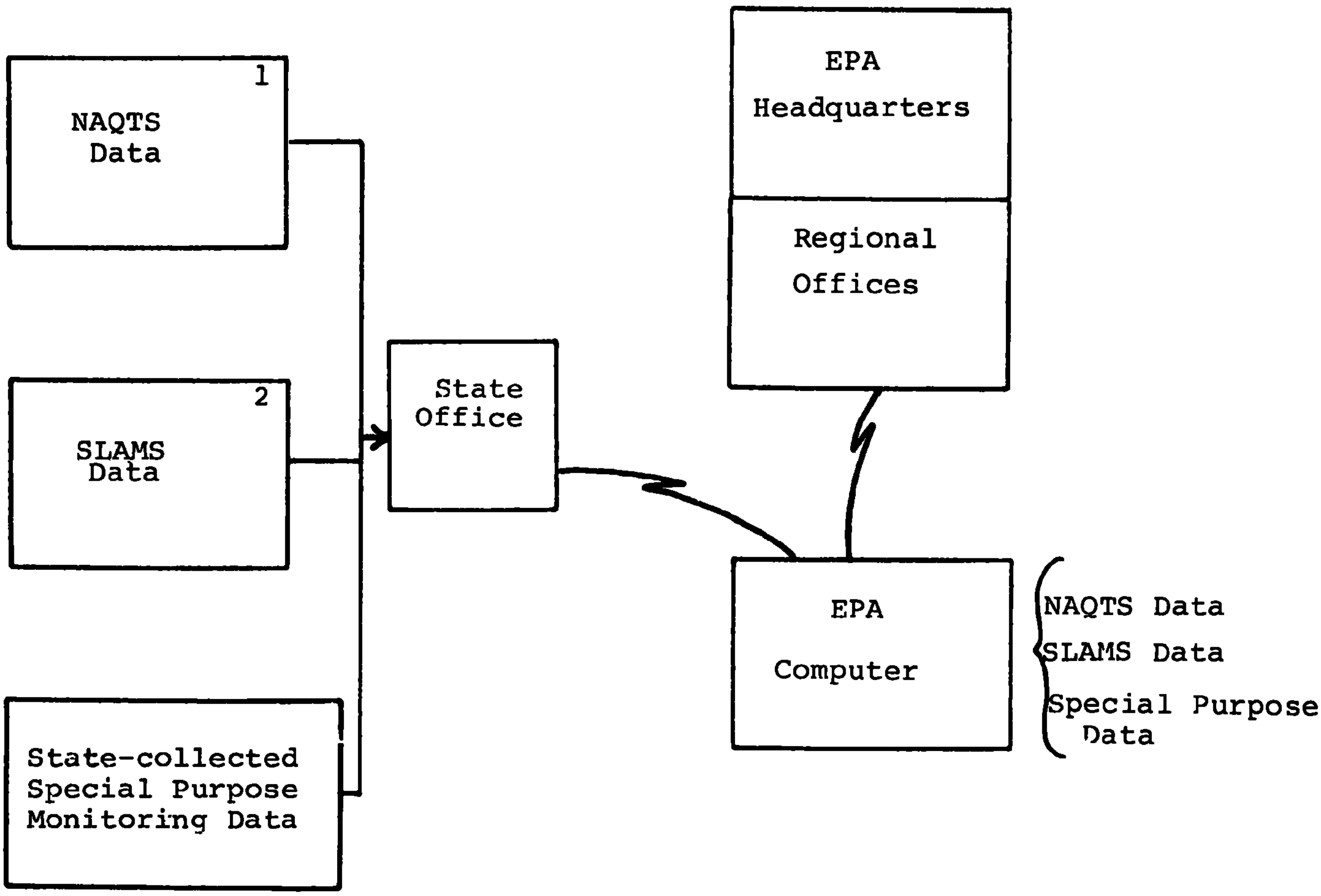
(1) Three Major Computer System Design Alternatives Could Be Used to Manage Air Quality Data.

EPA could adopt three different computer system design philosophies for managing air quality data, as follows;

- A utility approach in which states input all ambient data into a centralized EPA-operated data base
- A hybrid approach in which the bulk of data resides on state data files with a subset stored centrally on an EPA data base
- The current approach in which data are stored both on state data files and centrally on an EPA data base

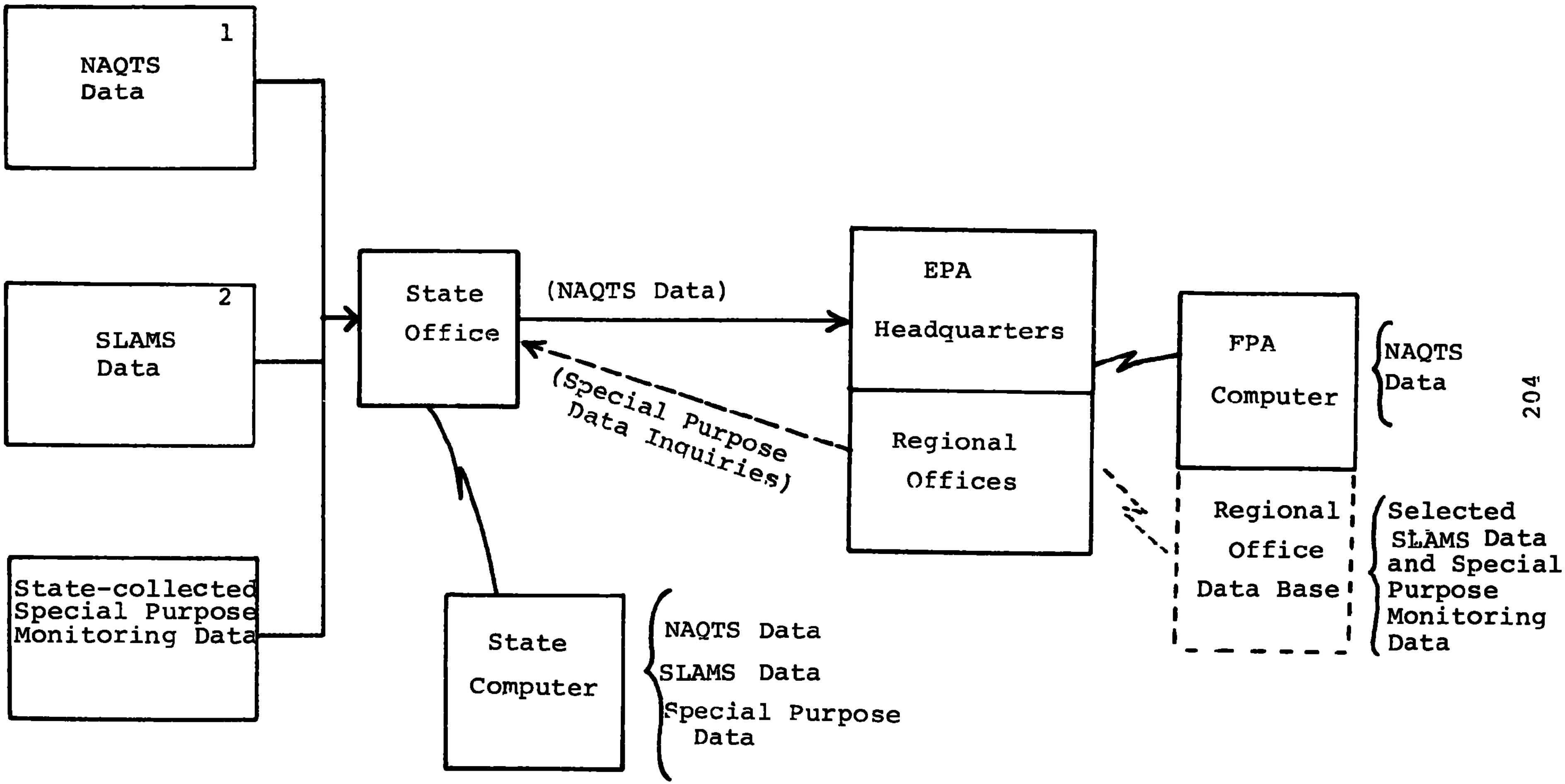
Exhibits 6-2 through 6-4, on the following pages, illustrate the flow of data between state offices and EPA for each alternative. Each approach represents a generic type of

Exhibit 6-2
U.S. Environmental Protection Agency
Utility Air Quality System Alternative



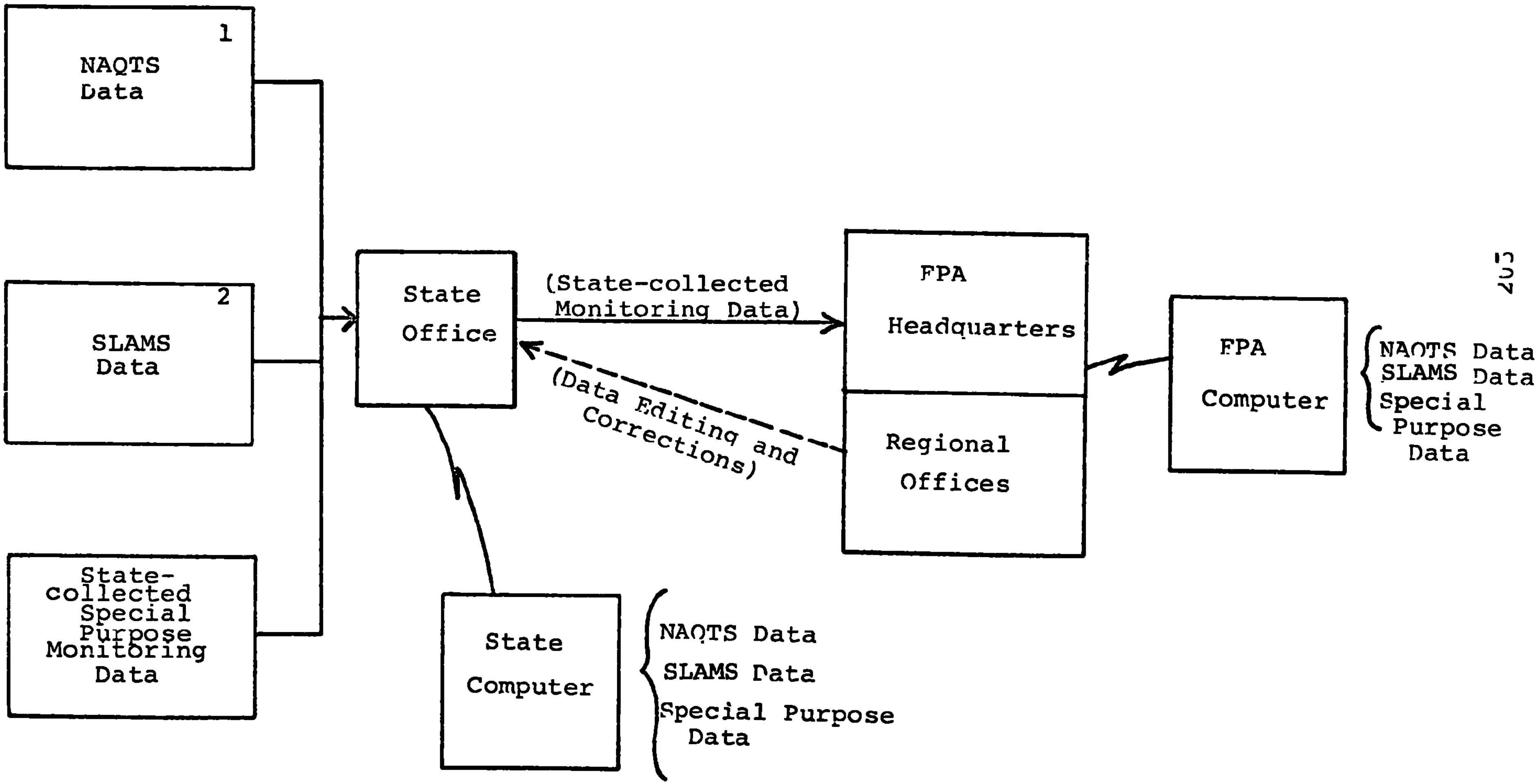
1 NAQTS - National Air Quality Trend Stations
2 SLAMS - State and Local Air Monitoring Stations

Exhibit 6-3
U.S. Environmental Protection Agency
Hybrid Air Quality System Alternative



1 NAQTS - National Air Quality Trend Stations
2 SLAMS - State and Local Air Monitoring Stations

Exhibit 6-4
U.S. Environmental Protection Agency
Current Air Quality System Alternative



1 NAQTS - National Air Quality Trend Stations
2 SLAMS - State and Local Air Monitoring Stations

system design. In turn, each generic type may have different technical variations involving minor differences in data flows, management and reporting responsibilities.

The utility alternative assumes that state offices would communicate to a central air quality data base at an EPA computing facility. All state-collected data, including State and National Air Quality Trend Station data (as defined by SAMWG) and special purpose monitoring data, would be maintained centrally. Both EPA and state offices would utilize the same data base for analysis. This approach is similar to the system design for STORET, EPA's storage and retrieval system for water quality data.

The hybrid alternative presumes that most air quality data would reside on state data bases. States would utilize their own air quality data systems to edit, update and analyze local data. A subset of this data, National Air Quality Trend Station (NAQTS) data, would be communicated to EPA and stored on a central data base for analysis. Those states which require air quality data systems support but could not operate their own system economically could transfer their data to EPA regions or Headquarters and could rely on regional or central systems for update and report retrieval of local data. Some Regional Offices may elect to operate a

CDHS system either on the central EPA computer or on a regional mini-computer. The regional systems could be used to provide automated support to smaller states or to provide supplementary information regionally as may be defined in the future by SAMWG. In most cases under the hybrid approach, Regional Offices and others requiring summary statistics or raw data from non-National Trend Monitoring Stations would obtain this information, as needed, from the state office which collects this special purpose data.

The current alternative is represented by the current SAROAD/AQDHS-II system design. State offices and EPA would continue to maintain separate systems operating at their respective computer facilities. All state-collected data would be transferred to EPA, periodically. EPA would utilize its own data base for analysis and would supplement its information, when necessary, by communicating with state offices.

(2) The Hybrid and Utility Design Approaches Provide Better Functional Service to EPA and State Offices Than Does the Current Approach.

Hybrid and utility system alternatives for air quality data would improve data handling and correction capabilities and would offer Headquarters and states flexible reporting and good data availability. Exhibit 6-5, on the following page, illustrates these advantages by comparing the functional service provided by the three alternatives to Headquarters, regional and state users.

The current design approach provides EPA and states with good data availability and reporting flexibility since each group uses its own data base for analysis. However, as a result of the large volumes of data transferred and the relative distance and processing steps between the sources of data and the EPA users, data handling and correction delays are created which contribute to completeness and currency problems of EPA-maintained data.

A utility design would improve the functional service to EPA. On the other hand, state service would be reduced with a utility approach because a large centralized data base would not allow for customized reporting to the same degree as would a state data base. Although a utility approach would eliminate date reconciliation problems because data would be entered only once, the size of the data base and associated system management problems would create additional operational problems.

Exhibit 6-5
U.S. Environmental Protection Agency
Functional Service Air Quality
Data System Design Alternatives

	Utility			Hybrid			Current		
	S	R	H	S	R	H	S	R	H
Reporting Flexibility	Fair	Good	Good	Very Good	Fair	Good	Very Good	Good	Good
Data Availability	Good	Good	Good	Very Good	Fair	Good	Very Good	Good	Good
Data Handling Editing and Correction	Good	Good	Good	Good	Good	Good	Good	Poor	Poor
Overall Functional Service	Improved data handling with good availability of state-collected air quality data			Improved data handling with good availability of monitoring trend data			Good data availability and reporting flexi- bility but with poor data handling		

S = State Offices
R = Retional Offices
H = Headquarters

The hybrid approach would improve data handling because less data would be transferred to EPA and fewer opportunities for errors would arise. Current SAMWG estimates indicate that 90% of the current reporting stations could be eliminated from the central EPA data base under the hybrid approach. States would be serviced equally well by current or hybrid designs since the states would maintain control over their own data bases and reporting. However, state-collected special study data would not be as available to regional offices under the hybrid approach. Regional offices would need to rely on states to supply data in support of various special programs, such as modeling applications.

The hybrid and utility approaches are comparable in functional service because each approach would service the states and Headquarters well while improving data handling and correction. Regional offices and others requiring non-national trend data would deal with data availability problems by communicating with state offices, as needed.

(3) The Hybrid Approach Is Slightly Less Expensive Than the Other Alternatives.

A hybrid approach for air quality data would reduce slightly aggregate computer operating costs for EPA and state offices yet would entail development and installation costs comparable to those necessary to support the current design philosophy.

Exhibit 6-6, on the following page, contrasts various cost considerations for the three alternatives. Appendix B presents more detailed derivations for the development and operating costs associated with each of these alternatives.

The combined computer operating costs to EPA and state offices with the current design approach are estimated to total $800,000. These estimates assume the completion of currently contracted modifications to reduce the operating costs of SAROAD. Included in the total cost estimate is the cost of additional data reporting that states provide to EPA because of EPA data correction problems resulting from reconciliation of the separate data bases. The system modifications and development work associated with AQDHS-II together with installation costs for additional states are estimated to total $330,000.

The computer operating costs for a utility design would be lower than those incurred with the current approach. However, the costs associated with developing and installing a user-oriented utility system would be $600,000 to $1,000,000. The utility approach would allow larger states to reduce manpower slightly but would require a slight increase in EPA staffing. Overall, comparable numbers of personnel would still be needed to handle data, initiate reports, coordinate and customize

Exhibit 6-6
U.S. Environmental Protection Agency
Cost Evaluation Air Quality
Data System Design Alternatives

	Utility	Hybrid	Current
Computer Operating Cost	$700,000	$600,000	$800,000
Development and Installation Cost	$ 600,000 - $1,000,000	$380,000	$330,000
Personnel Cost	Lower by approximately five people*	Same as present	Same as present

* This reduction could be translated into $75,000

system usage and oversee system update and operations. Net personnel savings in the states and EPA are estimated not to exceed five positions.

The computer operating costs for a hybrid approach are estimated to total $600,000, which is lower than the other two alternatives. The development and installation costs would be comparable to the current approach but lower than the utility approach. In addition, personnel costs would be similar to the levels experienced with the current design. However, these staffing levels would be sufficient to insure timely submission and correction of national trend station data. By contrast, these same staffing levels have proven inadequate to insure timely submission and correction of data under the current approach. Regional offices have indicated that as much as one extra person per region might be needed to insure the accuracy of current volumes of data.

The hybrid approach is perferable because combined EPA and state computer operating costs are lower than the other alternatives and because system development and installation costs are comparable to those that would be incurred under the current mode of operation.

(4) <u>The Hybrid Approach Is the Least Risky of the Three Alternatives</u>.

A hybrid system design would offer EPA lower technical and cost risks than the utility or current approach. Exhibit 6-7, on the following page, compares risk and timing considerations for the three alternatives.

The current approach presents a fairly high technical risk to EPA because of the large EPA data base size and the related UNIVAC system problems. The associated cost risk is moderate, but manageable, because EPA and states control their own systems., that is, states can decide for themselves how much special purpose storage and retrieval is cost-effective without requiring that EPA store all state-desired data as well.

On the other hand, the utility approach would present a very high technical and cost risk potential because many users would be dependent upon one large centralized system for all air quality data. Therefore, EPA would need to store all state-desired data whether or not these data were relevant for national reporting or analysis. Moreover, because of the service philosophy associated with the utility approach, data base size and operating costs would not be controlled easily.

Exhibit 6-7
U.S. Environmental Protection Agency
Air Quality Data System
Alternative Design Risks

	Utility	Hybrid	Current
Technical Risk	Highly sensitive to hardware problems	Less sensitive to hardware problems	Sensitive to hardware problems
Cost Risk	Data base size and operating costs not easily controlled	EPA and state control of respective operating costs	EPA and state control of respective operating costs
Timing	1-2 Years	2-3 Years	1-2 Years
Overall Risk	High technical and cost risk	Low technical and cost risk	Moderate technical risk

Conversely, the hybrid approach would present a lower technical risk to EPA because EPA would maintain a much smaller data base consistent with the SAMWG data reporting requirements. Similarly, the cost risk would be reduced since the majority of data processing would be controlled by the states who, in turn, would be free to decide and pay for whatever data they needed for local purposes. However, the hybrid approach may require more time to become operational than would the other alternatives because it presupposes a coordinated phasing of data system strategy with SAMWG monitoring strategy.

Overall, the hybrid approach is preferable because of its lower technical and cost risks.

(5) Management Considerations Favor the Hybrid Approach.

A hybrid design for air quality data would provide EPA with a system which is easy to manage and which would not entail heavy dependence of states on EPA-provided ADP support. Exhibit 6-8, on the following page, compares three management considerations for the feasible alternatives.

With a hybrid approach system management problems would be limited since EPA and states would maintain their own systems and only a small amount of data would be subject to reconciliation problems. However, EPA would be forced to rely

Exhibit 6-8
U.S. Environmental Protection Agency
Management Considerations of Air
Quality System Design Alternatives

	Utility	Hybrid	Current
EPA Dependence on States	EPA relatively independent of states	EPA dependent on states for special study data	EPA relatively independent of states
State Dependence on EPA	States highly dependent on EPA for ADP support	States independent of EPA	States independent of EPA
Ease of System Management	Very complex system which is difficult to manage	Easy to manage at EPA and state levels	Easy to manage at state level but sensitive to hardware problems at EPA
Overall Management Evaluation	High system management problems with state dependence on EPA for ADP support	Good ease of system management with EPA dependent on states for special study data	Low dependence of EPA and states upon each other

on states to supply non-national trend data, as needed, to support various special studies. On the other hand, state dependence upon EPA for ADP support would be maintained at a relatively low level. Furthermore, the hybrid approach would be consistent with the data reporting and handling requirement which SAMWG is developing currently.

The current approach presents more system management problems than a hybrid design would present because of the large EPA data base and associated data reconciliation problems. With the current design, EPA dependence upon state offices is better than the hybrid approach since EPA has its own data base to support special studies. Still, EPA would need to communicate with states to reconcile data and to acquire missing data, as at present. Under both the current and hybrid design approaches, states would not need to depend heavily on EPA systems support.

The utility approach would impose more difficult management problems. Management difficulties would increase as a natural result of the use of a large central data base to service a diverse set of users whose system objectives would tend to differ over time. Although a utility approach would reduce the dependence of EPA on state offices for additional data support, states would be much more dependent upon EPA for ADP resources than under the hybrid or current approaches. EPA managers have indicated that minimizing state ADP reliance on EPA is a very significant policy objective and that the federal-state political relationship can be serviced best by allowing the greatest state flexibility. This concern has resulted from state fears that Congressional budget pressures could force EPA to sever its continuing support for utility-type services to the states which, in turn, would impose serious information processing problems on state programs.

The hybrid approach is preferable because it entails the fewest system management problems and does not increase state dependence upon EPA for ADP support above current levels.

(6) <u>The Hybrid Approach for Air Quality Data Could Be Implemented by EPA Over a Two-to-Three-Year Period</u>.

The implementation of a hybrid system design for the management of air quality data would require phased action steps over a two-to-three-year period. This transition from the current approach to a hybrid approach would be coordinated with the SAMWG-proposed changes in monitoring strategy and could be accomplished through eight actions, as follows:

- Define the EPA air quality data reporting and handling requirements associated with the monitoring strategy developed by the Standing Air Monitoring Work Group (SAMWG)
- Complete modifications to AQDHS-II and proceed with state installations
- Identify several prototype states to test new data reporting requirements associated with the SAMWG study over a period of at least six months
- Define procedures for EPA regions and Headquarters to acquire detailed state-maintained non-national data, when needed
- Define summary and violations data needed by EPA for management analysis
- Conduct preliminary system design and prototype tests for ambient violations reporting
- Define procedures for EPA support for those states without data systems
- After prototype testing, develop a phased plan for other states to change data reporting to EPA over a one-to-two-year period

These actions would require planning and coordination between Headquarters, regional and state offices in order to insure that functional service would be maintained for system users during the system design transition period.

2. CHANGES IN DESIGN PHILOSOPHIES FOR EMISSIONS AND COMPLIANCE DATA SYSTEMS WOULD PRODUCE ONLY MARGINAL BENEFITS AT HIGHER COSTS THAN PRESENTLY INCURRED.

Major changes in other air data system designs are not cost-justifiable at this time. Only marginal benefits would be realized by changing the current system design approach for managing emissions data. Likewise, the integration of CDS and NEDS or the consolidation of air research data would produce no substantial benefits.

(1) Changing the System Philosophy for Managing Emissions Data Would Not Reduce ADP Expenditures.

Emissions data are used by EPA and state offices to support nine activities as follows:

- Revising and evaluating SIPS
- Reviewing new sources
- Preventing significant deterioration of air quality
- Maintaining air quality standards
- Running simulation models
- Developing and revising emissions factors
- Performing energy studies
- Establishing new source performance standards
- Developing national emission summaries

EPA managers have questioned whether a different system design, such as a utility or hybrid approach, would improve the completeness and currency of the emission data used to support these activities. In addition, preliminary SAMWG data reporting requirements indicate that small point source emitters could be eliminated from a centralized EPA data bank. Although these small sources are important to state and local agencies, SAMWG has indicated that EPA has little, if any, need for such data. Exhibit 6-9, on the following page, summarizes the advantages and disadvantages associated with the utility, hyprid and current approaches. For the purposes of analysis, the utility approach assumes that most point sources in the nation are stored centrally. The current approach allows for a phased elimination of smaller sources from the data base, whereas the hybrid approach assumes a massive reduction in the number of stored sources and source parameters. As indicated in the Exhibit, the current approach remains the most practical.

A hybrid system design does not represent a feasible alternative at this time since EPA emissions information requirements could not be serviced adequately if this approach were adopted. With a hybrid approach the majority of emissions data would reside on the state data bases. As a result, Headquaters would not have the necessary national emissions

Exhibit 6-9
U.S. Environmental Protection Agency
Summary of Emissions Data System
Design Approaches

	Utility	Hybrid	Current
Computer Operating Costs	$400,000	N.E.	$400,000
Development and Installation Cost	$ 600,000 - $1,000,000	N.E.	$250,000 - $430,000
Functional Service	Good data availability and improved data handling	Poor data availability for Headquaters programs	Good data availability and reporting feasibility
Risk	High technical and cost risk	Low technical and cost risk	Moderate technical risk
Management	System management problems and state dependence upon EPA for ADP support	Less complex system management	Low dependence of EPA and states upon each other

N.E. = Not estimated because the hybrid approach would not provide adequate functional support.

data available to support a variety of air planning and review activities. For example, the development and evaluation of new source performance standards and EPA's ability to analyze nationwide trends in industry emissions of specific pollutants would be hindered. Also, analyses associated with various energy studies could not be performed adequately without national emissions inventories. In addition, intra-regional activities, such as modeling applications that cross state geographic boundaries, would be hampered. For these data availability considerations, the hybrid approach is not justified currently as a feasible alternative.

A utility approach would provide for good data availability as does the current approach. However, a large centralized data base would not allow for customized reporting to the same degree as would separately controlled state data bases. Although the utility approach would eliminate data reconciliation problems between EPA and state data bases, the large centralized data base would present a high technical risk because of associated computing resource requirements and the susceptibility to ADP hardware failures.

The aggregate computer operating costs to EPA and state offices are estimated to total $400,000 for both the current and utility approaches. However, the cost risk potential is less for the current approach because EPA and states control their own system. The development and installation of a user-oriented utility system would be $600,000-$1,000,000 whereas system modifications, development work, and additional installation costs associated with the current NEDS - EIS/PR are estimated to total $250,000 - $430,000. Furthermore, by changing to a utility approach, no substantial personnel reductions would be realized. (Appendix B presents more detailed cost derivations for these alternatives).

Current and utility approaches present systems management difficulties. While the current design creates data reconciliation problems for EPA, a utility design would be more difficult to manage because of diverse user requirements. Also, states would be much more dependent upon EPA for ADP resources with a utility approach.

As a result of the increased cost potential and technical risk associated with the utility approach, the current system design should continue to be used to manage emissions data. However, EPA will need to coordinate any reductions in the scope of centrally-maintained point source data with future SAMWG recommended changes in emissions data reporting requirements.

(2) The Integration of CDS and NEDS Would Not Reduce ADP Expenditures.

Air program and enforcement managers have questioned whether emissions and enforcement data should be integrated into a single data base in order to improve reporting capabilities and reduce computer operating costs. They have suggested that the consolidation of the two systems would eliminate redundant data storage and would enable analysts to relate compliance status to emission abatement.

Although the two systems both contain data about stationary sources, less than 10% of the actual data values stored within the systems are common to both CDS and NEDS. Consequently, under an integrated system neither the volume of transactions nor the size of the data base would be reduced substantially. In addition, the sources of CDS and NEDS data differ so that consolidated data entry would not be feasible. Overall, an integrated system would not reduce operating costs but would require new systems development expenditures of at least $200,000.

Furthermore, the integration of data bases would not improve substantially the functional service to EPA, since most EPA personnel who need emissions data and those that need compliance information tend not to need both. Air program personnel use emissions data for planning, review and abatement activities, whereas air enforcement personnel need source compliance information to monitor conformity with SIP regulations. NEDS emission calculations are not current or precise enough to support enforcement actions. Rather, enforcement personnel use data acquired from source inspections, 114 inquiries and stack tests to support enforcement actions. Air program personnel use source compliance status data to support limited planning and review functions. NEDS currently has a facility for storing a one-character compliance status indicator which is sufficient for planning purposes.

In addition, the integration of CDS and NEDS would create substantial system management problems. Currently, CDS is supported by the Division of Stationary Source Enforcement (DSSE) within the Office of Enforcement and NEDS is supported by the National Air Data Branch (NADB) within the Office of Air Quality Planning and Standards. An integrated data base would be difficult to maintain because of the coordination and communication problems that would result from the different geographic locations and distinct organizational responsibilities of the system users. Also, the nature and timing of the compliance data flows differ from those associated with emissions information. These differences in data sources, timing and management responsibilities would cause serious disruption in functional priorities and services. The integration of CDS and NEDS would not improve system capabilities substantially nor would it reduce computer costs appreciably. Consequently, the current separate system approach remains the most practical.

3. TECHNICAL SYSTEM ACTIONS COULD IMPROVE FUNCTIONAL SERVICE AND SYSTEM EFFICIENCIES.

Various system modifications to improve data base design and to reduce computing resource requirements would make current data systems less susceptible to UNIVAC system failures. These changes would make the transfer of various air systems to another computing facility inappropriate and unnecessary. Additional system modifications to NEDS, SAROAD and CDHS, however, could streamline system operation further and, as a result, could improve system utilization.

(1) Additional SAROAD Modifications Could Streamline System Operation Still Further.

The current contractor modifications to SAROAD should improve system efficiency by at least 50% and, as a result, should improve data availability to the various system users. However, three additional SAROAD design modifications could further reduce processing time and make SAROAD more responsive to user needs, as follows:

- Modify the method for calculating running averages for certain pollutants
- Revise the timing of summary and frequency file creation
- Develop a violations file

The specific cost impact of each of these changes will depend heavily upon the results of current contractor enhancements. Whether or not to undertake these changes will depend therefore upon the resulting SAROAD efficiency and processing characteristics once the current enhancements are completed.

Current enhancements will cause running averages to be calculated from hourly monitoring data for SO_2, NO_2, and CO during the update of raw data. As a result, 38% of the values stored on the less-than-24-hour raw data file will be associated with running averages. However, storage costs could be reduced if running averages were not stored but were calculated and reported only when needed. Whether or not the frequency and distribution of running average requests would justify this design change cannot be ascertained from current NADB report logs. Instead, NADB should track reporting activity once the enhancements are completed to determine whether running averages need to be stored or could be recalculated when needed.

Under current modes of operation, summary and frequency files are created after each raw data update. However, the two processing steps do not have to be executed concurrently. Instead, computer operating costs could be reduced if the summary files were created at fixed time intervals over the year independently of raw data updates.

Various system users rely on SAROAD retrievals to provide information on air quality standards violations. Presently, this reporting is accomplished by comparing collected ambient data averaged over the appropriate time interval to the associated NAAQS. Computer processing time could be reduced, however, if a violations file were created which indicated all violations for a given site and time period. In this way, subsequent calculations would not be needed to reidentify violations for other analyses. Violations could still be calculated from raw data files, however, if ambient values were changed or if analysts wished to evaluate the violations impact of changed ambient standards.

None of these changes can be cost-justified at present with currently available information. Instead, NADB will need to analyze the costs and benefits of these proposed changes after current modifications have been completed.

Three areas could be analyzed in order to evaluate accurately these costs and benefits, as follows:

- Evaluate SAROAD modifications over a two-to-four period to ascertain achieved processing efficiencies

- Analyze the frequency of various retrievals by system user and determine the geographic distribution of retrieval requests
- Analyze the frequency of redundant violation calculations

This information is not available now because SAROAD modifications are still underway. However, it is anticipated that the frequency of retrievals will increase as the enhancements provide improved data availability.

(2) EPA and Emissions Inventory Information Support Could Be Improved Substantially Through Minor Enhancements to NEDS Edit Procedures.

Current enhancements to NEDS, including a regional edit capability and improved update efficiencies, will not be sufficient to improve emissions inventory accuracy and currency because regions still will have only limited editing capabilities. Many edit checks, and all validation checks, are scheduled to be performed during the NADB-initiated update and therefore cannot be executed by regions. As a result, regions still will have to wait two-to-three weeks to receive update listings and validation reports from NADB. These processing characteristics will continue to require extra regional effort to reconcile errors and will delay further the input of corrected NEDS data.

If regions had the ability to run validation checks and to edit input cards against the user file, then these problems would be reduced substantially. Current contractor efforts to modify the NEDS edit could be redirected to this end for only a minor additional cost since current enhancements have not yet been finalized and since the design specifications employ a modular approach, which could facilitate the redesign of the edits to run independently of the update.

This enhanced complete regional NEDS edit capability, similar to that offered by SAROAD, is necessary to maintain higher emissions inventory data quality and to facilitate regional processing efforts. This enhancement, in addition to the improved reporting capability which has recently been made available to regions, should improve significantly NEDS data accuracy problems. Approximately $5,000 - $10,000 in additional contractor funds are needed to give regions this enhanced edit capability.

It is anticipated that these changes should increase substantially the number of states submitting data to EPA on time. To date, some states have not been submitting data regularly because they felt that NEDS did not work. Furthermore, in some cases, regions have not discouraged this attitude because they were backlogged with data themselves and, as a result, were unable to input data quickly enough to maintain the data base currency. A workable system should improve state and regional reporting response.

(3) EPA Could Improve CDHS Performance by Proceeding With Five Identified System-related Actions.

Pressure for dependable CDHS support will increase substantially as EPA moves towards a hybrid approach for their management of ambient air quality data. Without enhancements, neither EIS/PR nor AQDHS-II would meet the data flow timing requirements of EPA and states as effectively as would be needed to support a hybrid systems approach. Rather, state and EPA users of hybrid systems would require better reporting, editing and tracking capabilities than are provided by CDHS at present. These features could be provided through five NADB-identified CDHS actions, as follows:

- Conduct review of states for both EIS/PR and AQDHS-II to identify additional system enhancements necessary to meet state information needs

- Complete current identified modifications to CDHS
- Perform additional enhancements to CDHS reporting capabilities as specified by state users
- Improve current data handling procedures for CDHS by incorporating more extensive edit capabilities
- Review and develop interface procedures between state-operated systems and EPA-maintained systems

The completion of currently identified CDHS modifications is necessary to improve the effectiveness of the system to state users. In addition, a joint EPA/contractor review of state users would identify additional areas for functional system enhancement. For example, many states have indicated that they would prefer EPA to develop more standard CDHS reports. A state review would provide EPA with a consolidated needs statement which would facilitate centralized enhancement efforts.

Improved data edit and validation procedures for CDHS would facilitate the transmittal of state data to EPA and would reduce the manpower required to correct data anomalies. Furthermore, improved CDHS/AEROS interfaces would eliminate the need to track data transmittals manually.

NADB has requested funding for these CDHS actions and has developed plans for their implementation during 1977. These

CDHS enhancements will improve the usefulness of state-operated data systems and thereby service more effectively EPA and state information requirements.

(4) The Transfer of Various Air Data Systems From the National Computing Center (NCC) to Another Computing Facility Is Not Cost-Justified at This Time.

EPA systems personnel have questioned whether the transfer of air data systems to a system environment more stable than the NCC would improve the usefulness of the systems. They have suggested that data processing operations could be performed in a more timely manner and that the accessibility of information would improve if data systems were operated on a facility that was less prone to frequent system crashes and prolonged downtime.

Although the NCC has experienced hardware and software problems since the beginning of the year, these difficulties are expected to be resolved and performance improved in the future. In addition, various system modifications are being implemented to improve data base design and to reduce computing resource requirements. These modifications should make the current data systems far less susceptible to computer failures.

In addition, the costs for running jobs at NCC will remain lower than the costs for running these programs at other EPA facilities such as OSI. EPA computer planning and operations personnel estimate that air systems costs at OSI would cost almost double the charge at NCC. In addition, since EPA owns the UNIVAC, the cost to the Agency for operating the UNIVAC would be incurred regardless of the computing load at NCC. As a result, overall agency ADP expenditures would increase if current NCC-operated systems were transferred to OSI or other computer facilities.

Should UNIVAC performance deteriorate once again, however, and should the dramatic reductions in AEROS processing time requirements not prove sufficient to allow adequate turnaround, then EPA will need to reconsider the possibility of moving AEROS to another computer. In particular, EPA should await the completion and testing for several months of current design changes before reconsidering such a move.

(5) A Management Audit of the Energy Data System Would Establish Whether the Current System Design Efficiently Fulfills Its Intended Purposes.

Energy Data System ADP costs have caused various EPA system managers to question the feasibility of the EDS system design and operation.

EDS was developed under contract to SASD in order to satisfy three needs, as follows:

- To provide an automatic means of correlating energy-related data from various sources
- To provide quicker turnaround of data analysis than would be available otherwise
- To provide an easier means of correcting erroneous data than is available for related systems

The energy data base currently maintains over 58 million characters on-line. The data base is increasing this year by approximately six million values. The data base size is expected to grow at a lesser rate in future years depending upon system manager and user needs. A large portion of these data are stored also in other AEROS systems. NADB managers have questioned whether the computing resources necessary to support the storage of this volume of redundant data are warranted. EDS support personnel estimate that FY 77 ADP costs will be $300,000. However, the original EDS feasibility study, which was used to authorize the system, had projected that ADP costs would average only $25,000 per year. However, the feasibility study did not anticipate that EDS would continue to require ADP expenditures for on-going enhancements and reporting. EDS managers have estimated that approximately $190,000 to $219,000 of these ADP costs are used for these enhancements. In addition, the structure of the data bases which supply information to EDS has changed since the original EDS feasibility study in 1974. As a result, EPA system managers have suggested that a different system design which directly accesses the principle source data bases may be more efficient than the current EDS design.

EDS system managers argue that because of the ad hoc nature of data requests, most system inquiries require a selective flexibly formatted reply within one day in order to be meaningful to decision-makers. Although this short response time and flexibility may be optimal, this rapid turnaround, in itself, may not cost-justify an interactive retrieval system. Furthermore, proposed modifications to the supplying AEROS data systems should improve their response time so that one-day turnaround might be possible from AEROS and thereby obviate the need for some of EDS's more expensive capabilities.

A common complaint expressed by various EDS users has been related to the completeness and currency of data contained in the supplying data systems. EDS system managers have expressed a need to update and to change data in EDS independently of changes in the other systems. Such modifications create data reconciliation problems between EDS and the supplying systems. Moreover, subsequent data transfers from emissions and air quality data bases could replace corrected data with erroneous data if the data base management between the affected systems were not coordinated.

A management audit is needed at this point in time to determine in what manner system operations should continue and to determine the level of ADP resources needed for continued EDS support. A management audit would answer five questions, as follows:

- For what purpose is EDS used? By whom? What are its information system requirements?
- Presently, how does EDS satisfy these requirements?
- How efficient is EDS as an information storage and retrieval system?
- Are the ADP costs of EDS justified in terms of decisions for which the system supplies data?
- Who should retain management responsibility for EDS?

This audit would enable EPA management to ascertain the value of the current system and to determine whether additional resources should be applied to it.

4. THE CLARIFICATION OF MANAGEMENT RESPONSIBILITIES WILL IMPROVE THE DEVELOPMENT AND USE OF AIR DATA SYSTEMS.

The development and operation of selected air data systems could be enhanced if EPA system managers would utilize more rigorous ADP management principles and would employ tighter controls on the use of computing resources.

(1) The Development and Operation of Air Data Systems Could Be Improved by Centralizing System Management Responsibilities Within OAQPS.

The management of system development efforts and data system operations requires full-time monitoring by personnel with a sound foundation in ADP management principles. This is necessary both to insure that data systems make efficient use of computing resources and to coordinate system development and operations between interrelated data systems.

The National Air Data Branch (NADB) is the organization within OAQPS with the widest expertise in system development and operation. NADB has been responsible for most of the data system development and operations associated with air program data systems. In those cases where NADB has not been primarily responsible for system management, such as with EDS, ADP management principles have not always been employed because of the unavailability of experienced ADP managers. In these instances, feasibility studies and design efforts have been undertaken without a full appreciation of the potential cost and complexity of seemingly simple system features. Only organizations with sufficiently broad and extensive experiences with multiple systems can be expected to have the level of expertise necessary to manage large development activities. For this reason, system management responsibilities should be centralized in OAQPS within NADB and should be administered by those personnel with the most experience and training in ADP management practices. However, in order to insure adequate service to the other divisions within OAQPS, NADB should receive formal direction from OAQPS senior management regarding office-wide priorities and support responsibilities.

(2) <u>A Detailed Technical and Operating Analysis of the UNIVAC and Its Management Is Needed to Improve Computer Reliability</u>.

NCC's recurring hardware failures have caused significant delays and complications for the air data system and other NCC users. A detailed analysis of hardware malfunctions, throughput volumes, channel capacities and operating system performance is needed to isolate specific areas for technical improvement. In addition, this analysis would examine the effectiveness of current NCC organization, management procedures, staffing levels and contractor utilization. As a result of such an analysis, EPA could determine whether or not to shift its current operating mix of batch and demand processing, to curtail future hardware changes or to revise its current data center management structure. This analysis is estimated to cost approximately $50,000.

(3) The More Widespread Utilization of Standard EPA ADP Policies and Procedures Would Force a Closer Alignment of System Capabilities with High Priority Program Needs.

In order to improve system effectiveness and performance, EPA system managers should adhere to EPA policies and procedures in three ADP related areas, as follows:

- Development and review of system feasibility studies
- Data system design and use
- Consistent documentation and training

Appropriate guidelines and procedures for these areas are discussed in the EPA Automated Data Processing Manual and other related documentation.

Although managers and technical personnel within NADB and DSSE have tended to conform to these principles, systems personnel in other parts of OAQPS and ORD have not always been as rigorous in their use of formal development controls. The contracting office in RTP has allowed some systems development projects to proceed without requiring formal and separate feasibility studies. For example, work is currently underway to enhance the CHAMPS data system using contractor personnel. This contractual effort was initiated without explicit design objectives or acceptance conditions established. Instead, the statement of work specified that the work effort would include both a feasibility analysis and development effort by the contractor.

EPA system managers often utilize contractors to develop feasibility studies for proposed data systems. However, conflicts of interest and objectivity problems arise when the same vendor is asked to assess the need for and to build a system. Accordingly, contracting officials should write and award feasibility study contracts independently of subsequent system design efforts so as not to compromise contractor efforts and to assure that an adequate cost-benefit analysis is performed. Furthermore, feasibility study preparation should be overseen and reviewed subsequently by EPA personnel with appropriate ADP management and technical expertise and not solely by program personnel with little or no systems background.

In addition, air data systems should make use of consistent documentation conventions and user training material. At present, air systems documentation and user training vary among the systems depending upon the standards employed by the original system developers. NDB has established documentation conventions to be used for all systems operated under its jurisdiction. These same guidelines should be applied to systems operated within ORD and to those CDHS systems provided to states.

(4) Tighter Controls on the Use of ADP Resources Need to Be Employed by EPA System Managers.

As a result of Agency-wide ADP budget constraints, tighter controls on the use of ADP resources will need to be employed by EPA system managers. Since a plan has been prepared to develop the NCC into an Agency-wide resource which will be priced competitively with the IBM computer at OSI, MIDSD anticipates that the competition for UNIVAC resources will increase. As a result, not all air program computing requests will be satisfied. Consequently, EPA system managers will need to analyze their ADP utilization to insure that current and future applications operate efficiently in order to free enough system resources to preserve sufficient funds for high priority or emergency projects.

Better computing accounting records than have been maintained currently would provide system managers with more detailed information on ADP utilization. System managers should establish ADP project codes to reflect the development, maintenance and reporting costs associated with the various data systems and system users. With these sub-categories, system managers could ascertain the ADP costs associated with data reporting for program offices other than their own. Furthermore, this cost information could provide MIDSD with a basis with which to reallocate portions of various ADP sub-allowances to reflect the data reporting that one program office performs on behalf of another.

* * * * * * * * * * * * * * *

In summary, over the next two-to-three years a fundamental change in the data system philosophy employed by EPA will be needed to manage air quality data. A hybrid system design approach will improve service to EPA while reducing associated operating costs. On the other hand, comparable design changes to other systems will not be needed. Rather, technical system changes will suffice in improving system efficiencies and usefulness and thereby will minimize many of the data handling and reporting problems which have been experienced to date. These changes will require also certain shifts in ADP resource allocations and management responsibilities. Overall, air data system design and utilization need not undergo major additional changes, but rather can evolve in a more gradual manner to satisfy changing functional requirements.

VII. RECOMMENDATIONS

VII. RECOMMENDATIONS

This chapter presents the recommended actions which EPA should perform in order to improve the effectiveness and functional adequacy of the air data systems. The recommendations are grouped into three categories, as follows:

- Actions to implement a hybrid system for the management of air quality data
- Technical system actions to improve air data system efficiency and utilization
- EPA management actions to improve ADP resource utilization

In addition, this chapter presents the projected ADP expenditures, contractor funding and EPA staffing associated with the air data systems over the next two years. The cost estimates for the proposed actions in this chapter are summarized in Appendix B.

1. EPA SHOULD INITIATE FIVE ACTION STEPS OVER THE NEXT TWO-TO-THREE YEARS IN ORDER TO ESTABLISH A HYBRID SYSTEM FOR THE MANAGEMENT OF AIR QUALITY DATA.

EPA should adopt a hybrid system approach for managing air quality data. The implementation of this hybrid system should be phased over the next two-to-three years to facilitate management control and to insure a smooth and non-disruptive phasing of individual states into the hybrid reporting system.

The implementation process should be divided into five steps, as follows:

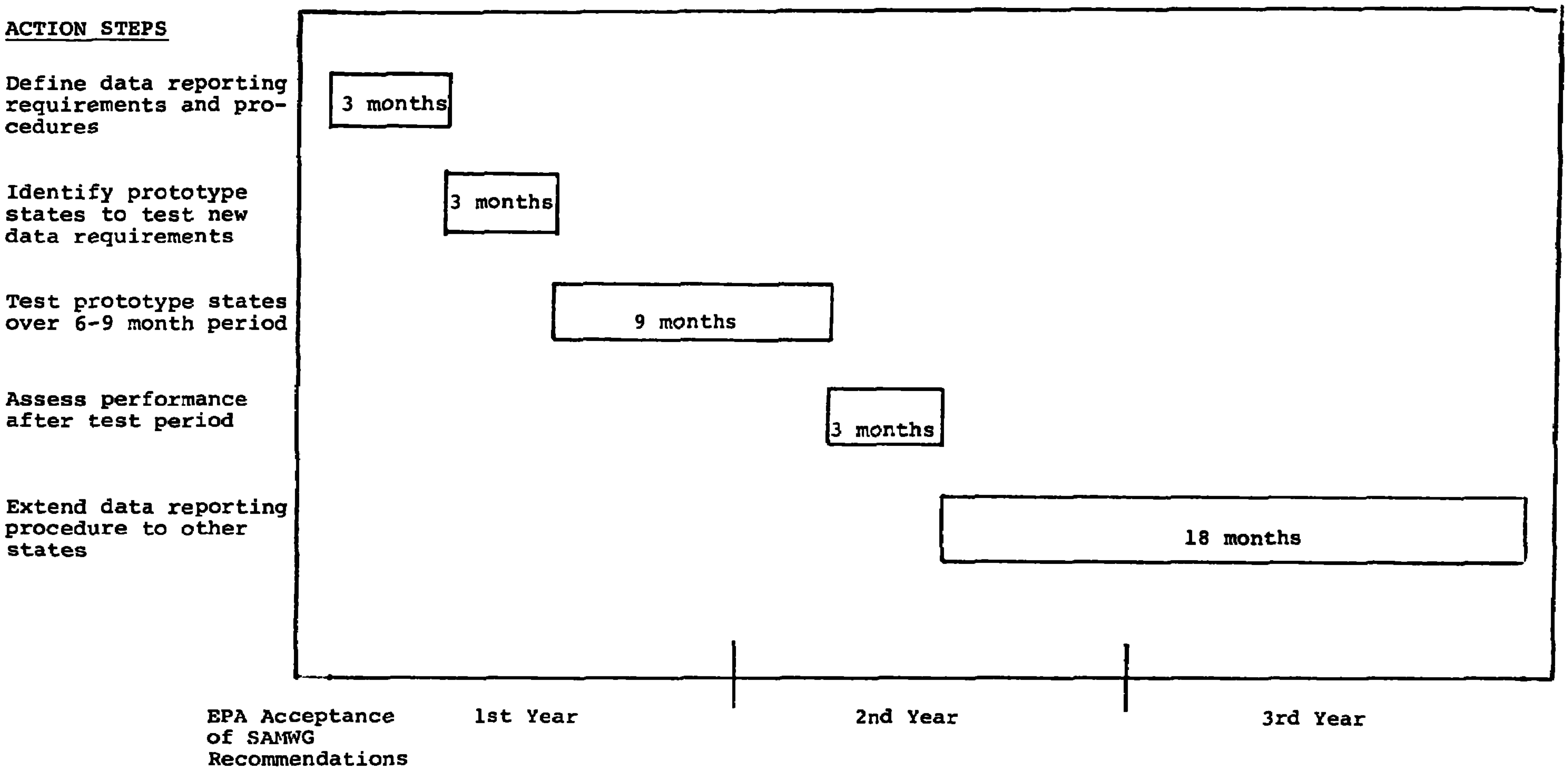
- Define data reporting requirements, procedures and control agency grant implications
- Identify prototype states to test new data requirements
- Test prototype states over a six-to-nine month period
- Assess performance after test period
- Extend data reporting procedures to other states

Exhibit 7-1, on the following page, illustrates the time-phasing for each action step. The segmentation of the implementation process into well-defined phases will facilitate control by providing checkpoints whereby management can evaluate the progress of each phase and initiate corrective actions where necessary.

(1) EPA Should Define the Specific Date Reporting Requirements and Procedures Needed to Support the Air Monitoring Strategy Developed by the Standing Air Monitoring Work Group (SAMWG).

OAQPS should develop a data system plan to support the EPA air monitoring strategy and should assume

Exhibit 7-1
U.S. Environmental Protection Agency
Time-Phased Actions Step for Implementing the Hybrid Air Quality System



responsibility for its phased implementation. The first step in this plan should be the definition of the specific data reporting requirements and procedures needed to support the monitoring strategy. These procedure and requirements definitions should include seven items, as follows:

- Timing, mode and format of the raw national trend station data to be submitted from state offices to OAQPS
- Specific summary and violations information needed from states
- EPA system requirements to support management analysis of state summary and violations information
- Plans for additional installations of AQDHS-II
- Procedures to interface with state-developed data systems other than AQDHS-II
- Procedures for EPA Headquarters and regional offices to acquire non-national trend station data from states
- Procedures to support those states without air quality data systems either with regionally-operated systems or with manual systems

These definitions can be formulated by OAQPS and regional personnel should require approximately three months' elapsed time after the publication of the air monitoring strategy by SAMWG. Alternatively, this analysis could be performed by outside contractors for approximately $25,000 during the same elapsed time period. In addition, EPA should examine the implications of these changes on regulations and control agency grant policies and revise these policies as deemed appropriate.

(2) OAQPS Should Coordinate With Regional Offices the Identification of Several Prototype States to Test the Monitoring Strategy Data Requirements.

Three states should be identified by EPA to serve as prototypes in testing the data reporting requirements and procedures associated with the EPA air monitoring strategy. These states should be selected to reflect four criteria classes, as follows:

- Differing industrial and residential mixes
- Different geographical and environmental regions of the country
- Varying population distributions
- Varying local data handling and analysis capabilities

The selection of these states should be a coordinated effort between OAQPS, regional offices and interested states. This identification and selection process should be completed within three months after the data reporting requirements are defined. During this same time period, EPA should design and develop a management summary and violations data system to process the summary data to be submitted by the state. This management summary and violations data system is estimated to cost $50,000 to design and implement and may be developed from existing SAROAD reporting modules.

(3) <u>The New Data Reporting Procedures Should Be Tested for the Selected Prototype States Over a Six-to Nine Month Period.</u>

The data reporting procedures should be tested for the selected states over two or three reporting periods so as to provide enough operating experience to assess performance. In order to coordinate activities during the test period, senior management within OAQPS, regional and state offices must commit adequate resources for support. Index estimates that one person within OAQPS and one-half person in each associated region and state office will be needed for the duration of the test period to coordinate activities.

(4) <u>After the Test Period EPA Should Assess the Performance of the Data Reporting Procedures</u>.

After approximately nine months of operation EPA should assess the performance of the data reporting procedures in the prototype states. The performance should be evaluated on three criteria, as follows:

- Accessibility of data to users in support of their respective program objectives
- Cost history
- Ease of management

With the assistance of the state offices, EPA should identify problem areas and should develop solutions before extending the procedures to additional states. Should special problems arise, it may be appropriate to select additional prototype states and to test for an extended period of time.

(5) <u>EPA Should Extend the Data Reporting Procedures to Other States Over an 18-Month Period.</u>

After the problems associated with the new data reporting have been identified and solutions have been developed, EPA should extend the distributed approach HYBRID to other states in phased steps. To insure control, EPA should rank states in terms of risk and priority, and should implement the hybrid approach first in those states which would be least susceptible to a change in their data handling interaction with EPA. In most cases, regional CDHS coordinators will take responsibility for assisting the states in the reporting changeover. The 18-month implementation period should be segmented so as to provide interim checkpoints which would allow for periodic progress evaluation. By the end of the implementation period all additional AQDHS-II installations should be completed. Furthermore, EPA should provide data processing support to those states that require, but can not support economically, automated data handling capabilities.

2. EPA SHOULD PERFORM FOUR TECHNICAL SYSTEM ACTIONS OVER THE NEXT SIX MONTHS TO IMPROVE AIR DATA SYSTEM EFFICIENCY AND EFFECTIVENESS.

The short-term changes to NEDS and SAROAD should be limited to those currently contracted including the use of contractor funds to implement CDHS enhancements. Before allocating additional resources for EDS support and system enhancement, however, EPA should conduct a system audit to reconfirm the need for proposed new and expensive current features.

(1) EPA Should Conduct Immediately a System Audit of EDS Before Assigning Additional Resources for Support and Development.

EPA should conduct immediately a system audit of EDS to determine the value of the current system and to ascertain whether additional resources should be applied to support it. The statement of work for this audit should include five areas for investigation, as follows:

- The air program information requirements to be supported by the system objectives
- The adequacy of the current system in satisfying these requirements
- The efficiency of the current system design
- Operating cost analyses in terms of the decisions for which the system provides information
- System management responsibilities

The estimated cost for the EDS audit is $15,000 and the elapsed time is estimated to be three months.

(2) <u>Short-Term Changes to SAROAD Should Be Limited to Those Currently Being Performed by Outside Contractors</u>.

Modifications to SAROAD should be limited to those currently underway by contractors. Long-term changes should not be authorized until a plan is developed by OAQPS for the implementation of the hybrid approach for managing air quality data and until NADB assesses the performance of the current enhancements under a stable AEROS software environment.

(3) Modifications to NEDS Should Be Limited to Those Currently Contracted Except That Data Edits Should Be Made Available to Regional Offices.

Enhancements to NEDS should be limited to those currently contracted, including a regional edit capability and improved update efficiencies. Furthermore, the current contractor efforts to modify the NEDS edit should be redirected to provide the regions with the ability to run validation checks and to edit input cards against the user file. Approximately $10,000 in additional contractor funds are estimated to be needed for this effort.

(4) EPA Should Use Its Budgeted $250,000 in Contractor Funds for CDHS Enhancements.

EPA should maintain its support for CDHS by taking five system-related actions, as follows:

- Complete current identified modifications to CDHS
- Conduct review of states to identify additional CDHS enhancements necessary to meet state information requirements
- Perform additional enhancements to CDHS reporting capabilities as specified by state users
- Improve current data handling procedures for CDHS by incorporating more extensive edit capabilities
- Review and develop interface procedures between state-operated systems and EPA maintained systems

NADB has estimated that these enhancements will require $250,000 in funds included in the FY 77 budget. The elapsed time for these actions is estimated to be approximately 18 months.

3. AFTER CURRENT AEROS SYSTEM MODIFICATIONS ARE COMPLETED EPA SHOULD EVALUATE SYSTEM PERFORMANCE BEFORE AUTHORIZING ADDITIONAL ENHANCEMENT EFFORTS.

EPA should evaluate AEROS system performance after current system modifications have been completed. This assessment should be performed over a three-month period to evaluate system efficiency and functional adequacy in a stable software environment. After this test period, NADB should analyze four areas for further modification, as follows:

- Consider alternative methods for calculating running averages for ambient data
- Revise the timing of summary and frequency file creation
- Analyze the need for a violations file for ambient data
- Consider altering the scope of source emissions data stored in NEDS consistent with any SAMWG recommended change in data reporting requirements

This analysis is necessary to ascertain costs and benefits of additional changes in light of the direction of future air program data requirements. This analysis can be performed by NADB personnel and should not require outside contractor assistance.

4. EPA SHOULD TAKE FOUR MANAGEMENT ACTIONS TO IMPROVE ADP RESOURCE UTILIZATION.

System development and operation responsibilities should be be centralized within OAQPS. MIDSD should conduct a detailed analysis of the UNIVAC system problems. In addition, MIDSD should encourage the use of standard ADP management policies, especially within ORD, and should provide guidelines to system managers on ADP accounting procedures and cost control.

(1) OAQPS Should Centralize System Operation and Development Responsibilities.

System management and development responsibilities should be centralized in OAQPS within NADB. Furthermore, in order to insure adequate service to the other divisions within OAQPS, NADB should receive formal direction from senior OAQPS management regarding office-wide priorities. The establishment of formal office-wide rather than divisional priorities should eliminate the need for the other divisions within OAQPS to develop new ADP systems on their own.

(2) <u>NEDS and CDS Should Not Be Integrated and Separate System Management Responsibilities Should Be Retained by OAQPS and DSSE</u>.

The current separate system approach for the management of compliance and emissions data should be maintained. Correspondingly, the system management responsibilities for NEDS and CDS should be retained by OAQPS and DSSE, respectively, as at present. In this way each organization will continue to control its own data base in a manner which is both practical and manageable.

(3) <u>MIDSD Should Conduct a Detailed Technical and Operating Analysis of the UNIVAC and Its Management.</u>

MIDSD should conduct a detailed technical analysis of the UNIVAC in four areas, as follows.

- Hardware malfunction
- Throughput volumes
- Channel capacities
- Operating system performance

This study should include an analysis of NCC processing characteristics and performance with other comparable UNIVAC installations. In addition, this study should analyze the adequacy of current staffing levels, management procedures, use of contractors and organization. As a result of this study MIDSD should develop an organizational and operations plan to reduce the overall failure rate at NCC. This analysis is estimated to cost $50,000 and could be completed within four months.

(4) MIDSD Should Encourage the Use of Formal ADP Management Policies and Procedures.

ADP policies and procedures are discussed in the EPA Automated Data Processing Manual and other related documentation. MIDSD should encourage the use of these procedures in four areas, as follows:

- Development and review of system feasibility studies
- Data system design conventions and utilization
- System and user documentation and training
- Development of formal ADP budgets reflecting full anticipated computer usage

The adherence by EPA system managers to these procedures will reduce the likelihood of developing inefficient data systems whose usefulness may be outweighed by their costs. The Office of Research and Development, in particular, should be encouraged to adopt these standard ADP management control procedures. Correspondingly, senior ORD managers and their technical ADP support personnel should stop treating ADP budgets as if they are interchangeable with other research funds and should begin to manage their ADP sub-allowances more rigorously so as to insure more effective use of computing resources.

5. <u>ADP EXPENDITURES IN SUPPORT OF THE AIR PROGRAM SHOULD REMAIN RELATIVELY STABLE OVER THE NEXT TWO YEARS</u>.

As a result of technical efficiencies, combined contractor and ADP costs will not increase significantly compared to current expenditure levels. Over the next two years annual ADP expenditures will remain relatively constant while contractor costs will decrease by an average of $250,000 annually.

Exhibit 7-2, on the following page, compares FY '76 estimated ADP expenditures by organization with the projected future annual costs. OAQPS will experience the most significant reduction in computer operating costs of approximately 40% compared to FY '76 expenditure levels. The Office of Enforcement has estimated that ADP expenditures will decrease slightly over the next two years. On the other hand, regional and state ADP costs will increase by approximately $375,000 as additional CDHS systems are installed and as local and regional personnel assume more responsibility for data analysis and modeling activities. Index has not been able to obtain projected ADP operating costs from either ORD Headquarters or laboratory personnel. However, we have estimated ORD operating costs to increase by 20% annually to reflect new system development and continued research analysis.

Exhibit 7-2
U.S. Environmental Protection Agency
Projected ADP Operating Costs by
Organization

	1976 ($000)	1977 ($000)	1978 ($000)
OAQPS	1656	1360	1066
ORD	857	1000	1200
OE	54	50	50
Regions	300	300	400
States	300	380	475
TOTAL	3126	3090	3190

Exhibit 7-3, on the following page, compares FY '76 estimated expenditures by system with projected future annual costs. The most significant reduction in operating costs will result from current technical SAROAD modifications. These technical changes plus the shift towards a hybrid approach for managing air quality data will reduce SAROAD computer costs by approximately 67%. On the other hand, CDHS costs will increase significantly as additional installations occur. EPA system managers have projected EDS operating costs to increase to $300,000 next year. However, we anticipate that an EDS system audit will identify areas where operating costs could be reduced back to current-year levels. ORD personnel were unable to provide Index with projected system operating costs. However, Index estimates that RAPS and CHESS/CHAMPS costs will increase by approximately 20% annually while system costs associated with the NASN program will remain stable.

Overall ADP contractor costs will decrease over the next two years. The table below compares current EPA contractor expenditures for OAQPS, OE and state offices with those anticipated in future years.

	1976 ($000)	1977 ($000)	1978 ($000)
OAQPS	800	750	550
OE	75	60	60
States	510	205	205

Exhibit 7-3
U.S. Environmental Protection Agency
Projected ADP Operating Costs by System

	1976 ($000)	1977 ($000)	1978 ($000)
SAROAD	903	480	300
NEDS	150	180	180
CDS	105	100	100
RAPS	250	300	360
CHESS/CHAMPS	347	410	490
EDS	183	300	186
CDHS	300	381	475
NASN	54	54	54
Other OAQPS Systems	380	400	400
Other ORD Systems	200	230	290
Other (including unallocated regional usage)	254	255	355
TOTAL	3126	3090	3190

Most ADP contractor expenditures were estimated from detailed contract listings contained in the Contracts Information System (CIS) and adjusted, where necessary, by system manager estimates.

OAQPS contractor expenditures for air data systems will decrease over the next two years. SAROAD and NEDS contractor costs will be offset by budgeted increases in CDHS support. While EDS contractor costs are expected to continue through FY '77, they are expected to be eliminated in FY'78 as a result of the proposed audit. In addition, approximately $75,000 in contractor support may be needed to implement the hybrid approach for air quality data. The Office of Enforcement cost has been estimated to decrease slightly in the future. State contractor costs will decrease slightly over the next two years. Index has not been able to obtain ADP contractor costs from either ORD Headquarters or laboratory personnel. Personnel levels in support of the air data systems are expected to remain stable during the next two years at Headquarters, in the regions and in the states.

APPENDIX A - INVENTORY OF CURRENT SYSTEMS

APPENDIX A - INVENTORY OF CURRENT SYSTEMS

The major ADP systems currently used by EPA and state air program personnel are described in this appendix. Each system write-up contains three types of information, as follows:

- System description
- System costs
- System management

Information contained in this appendix was obtained through interviews with EPA system managers and through the reviews of various system documentation.

STORAGE AND RETRIEVAL OF AEROMETRIC DATA (SAROAD)

SYSTEM DESCRIPTION	
Computer	UNIVAC 1110
Facility	NCC
Update Frequency	Scheduled bi-weekly but, in practice, irregular.
Data Base Accessible On-Line	Yes
Number of Standard Reports	32
Year Installed	1974-1975
Data Entry Transactions per Month	170,000
Transaction Size (Average)	6-9 data elements/transaction
Input Frequency	Continuous
Data Editing	Pre-edit uses no master files Edit uses two files
Programming Language for Update	COBOL, FORTRAN
Estimated Program Size	104 programs
Data Retrieval Method Used	FORTRAN, COBOL
Master File Size	88,000,000 values
Retention Period for Master Files	After three years old data stored on historical file
Data Base Management System Used	None
SYSTEM COSTS	
Fiscal 75 Computer costs	$ 606,242
Fiscal 76 Computer Costs	$ 903,000

SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	No
Contractors for System Maintenance	No
User Training Manuals	Yes

NATIONAL EMISSIONS DATA SYSTEM (NEDS)

SYSTEM DESCRIPTION	
Users	OAQPS, regional offices, states, other EPA, outside agencies and private institutions
Computer	UNIVAC 1110
Facility	NCC
Year Installed	1975
Update Frequency	Bi-weekly
Data Base Accessible on-line	No, although reports can be requested interactively
Number of Standard Reports	39
Data Entry Method	Batch
Software Package Used	None
Average Number of Transactions/Month	30,000-40,000 cards
Transaction Size (Average)	5-8 values per card
Entry Input Schedule	Continuous
Programming Language for Update	COBOL
Estimated Program Size	57 COBOL, 23 FORTRAN
Data Retrieval Method Used	COBOL, FORTRAN
Number of Master File Records	93,200
Master File Record Size	552 characters
Retention Period for Master Files	Indefinite
Data Base Management System Used	None
SYSTEM COSTS	
Fiscal 75 Computer Costs	$260,000
Fiscal 76 Computer Costs	$150,000

SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	No
Contractors for System Maintenance	No
User Training Manuals	Yes

COMPLIANCE DATA SYSTEM (CDS)

SYSTEM DESCRIPTION	
Users	DSSE, regional and state air enforcement personnel
Computer	IBM/370
Facility	OSI
Year Installed	1973
Update Frequency	Weekly
Data Base Accessible On-Line	No
Number of Standard Reports	4
Method	Batch
Software Package Used	None
Transactions per Month	40,000
Data Entry Input Frequency	Continuous
Programming Language for Update	COBOL
Data Retrieval Method Used	COBOL
Number of Master File Records	30,000
Retention Period for Master Files	Permanent
Average Number of Transactions per Update	10,000
Data Base Management System Used	None

COMPLIANCE DATA SYSTEM (CDS)

SYSTEM COSTS	
Fiscal 75 Computer Costs	\$50,000 (DSSE) \$30,000 (Regions)
Fiscal 76 Computer Costs	\$54,000 (DSSE) \$51,000 (Regions)
Fiscal 75 Contractor Costs(DSSE)	\$75,000
Fiscal 76 Contractor Costs(DSSE)	\$75,000
Current Manpower Support	1 (DSSE) 5 (Regions)
SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	Yes
Contractors for System Maintenance	Yes
User Training Manuals	Yes

ENERGY DATA SYSTEM (EDS)

SYSTEM DESCRIPTION	
Users	SASD
Computer	UNIVAC 1110
Facility	NCC
Year Installed	1975
Data Base Accessible On Line	Yes
Update Frequency	Yearly or quarterly depending on data values, or as needed
Average Record Size	N/A
Number of Data Values	58,000,000
Annual Data Base Increase	6,000,000
Data Base Management System	System 2000
Programming Language	COBOL, FORTRAN
SYSTEM COSTS	
1975 Computer Costs	$ 18,000
1976 Computer Costs	$183,000
1974-75 Contractor Obligations	$273,000
1976 Contractor Obligations	$129,000
Current Manpower Support	2 man years
SYSTEM MANAGEMENT	
Management	SASD and NADB
System Documentation	Not complete
Documentation - update procedures	Not complete
Contractors for system operation	No
Contractors for system maintenance	Yes
User documentation	Yes
EPA support personnel	2+

EMISSIONS INVENTORY SYSTEM (EIS)

SYSTEM DESCRIPTION	
Users	States
Computer	Varies
Facility	State facilities
Year Installed	1973
Update Frequency	Any time
Data Base Accessible on-line	Batch only
Number of Standard Reports	17
Number of Systems in Operation	14
Data Entry Method	Batch
Transactions per Month	500
Transaction Size (Average)	3,000
Input Frequency	Varies
Editing Procedure	Edit occurs during update
Programming Language for Updates	COBOL
Estimated Program Size	16
Data Retrieval Method Used	COBOL
Number of Master File Records	45,000-300,000
Master File Record Size	500-1,400 depending on number of pollutants, permits and registrations
Retention Period for Master Files	Permanent
SYSTEM COSTS	
Initial Development Costs	$150,000
Implementation Costs	$17,000/each
Fiscal 75 Computer Costs	$9,000 annually per state
Fiscal 76 Computer Costs	$9,000 annually per state

EMISSIONS INVENTORY SYSTEM (EIS)

SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	No
Contractors for System Maintenance	Yes
User Training Manuals	Yes
Number of EPA Personnel in System Management and Operations	1
Number of State Personnel in System Management and Operations	1-2 per state

AIR QUALITY DATA HANDLING SYSTEM (AQDHS-II)

SYSTEM DESCRIPTION	
Users	States
Computer	Varies
Facility	State facilities
Year First Installed	1973
Number of Systems in Operation	11
Update Frequency	Depends on state
Data Base On-Line Accessibility	No
Number of Standard Reports	4
Data Entry Method	Batch
Software Package Used	No
Transactions per Month	4,000
Transaction Size (Average)	6-8 data elements/transaction
Input Frequency	Varies
Data Editing Procedures	Uses site and parameter files
DATA MANAGEMENT	
Programming Language for Update	COBOL, FORTRAN
Estimated Program Size	15
Data Retrieval Method Used	Report retrieval package
Number of Master File Records	Varies between states (40,000-500,000)
Master File Record Size	Variable length - 41-133
Retention Period for Master Files	Indefinite
Average Number of Transactions per Update	Varies
Data Base Management System Used	None

AIR QUALITY DATA HANDLING SYSTEM (AQDHS-II)

SYSTEM COSTS	
Initial Development Costs	$150,000
Implementation Costs	$17,000 per system
Fiscal 75 Computer Costs	$9,000 per state
Fiscal 76 Computer Costs	$9,000 per state
SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	No
Contractors for System Maintenance	Yes
User Training Manuals	Yes
Number of EPA Personnel in System Management and Operations	1
Number of State Personnel in System Management and Operation	1-2 per state

SOURCE TEST DATA SYSTEM (SOTDAT)

SYSTEM DESCRIPTION	
Year Installed	1976
Computer	UNIVAC
Facility	NCC
Update Frequency	Variable
SYSTEM COSTS	
Fiscal 76 Computer Costs	$12,000
Fiscal 74-75 Contractor Costs	$34,442
Fiscal 76 Contractor Costs	$62,653
Manyears	½-1 manyear

STATE IMPLEMENTATION PLANNING SYSTEM (SIPS)

SYSTEM DESCRIPTION	
Users	Regions, OAQPS
Computer	UNIVAC 1110
Facility	NCC
Year Installed	1975
Update Frequency	Variable
SYSTEM COSTS	
Fiscal 75 Computer Costs	$ 2,426
Fiscal 76 Computer Costs	$ 9,600
Fiscal 75 Contractor Costs	$24,690
Fiscal 76 Contractor Costs	$38,000
Current Manpower	1 1/2 manyears

THE HAZARDOUS AND TRACE SUBSTANCE INVENTORY SYSTEM (HATREMS)

SYSTEM DESCRIPTION	
Computer	UNIVAC 1110
Facility	NCC
Year Installed	Projected for fiscal year 77
Update Frequency	Unknown
SYSTEM COSTS	
Fiscal 76 Computer Costs	Not yet implemented
Fiscal 75 Contractor Costs	$25,000
Fiscal 76 Contractor Costs	$50,000
Current Manpower	½ manyear
Expected Manpower	1-2 manyears

REGIONAL AIR POLLUTION STUDY (RAPS)

SYSTEM DESCRIPTION	
Users	ORD
Computer	UNIVAC
Facility	NCC
Year Installed	1973
Update Frequency	Every ten days
Data Base Accessible On-Line	Yes
Number of Standard Reports	10, plus numerous special reports
Storage Medium	Tape
Software Package Used	None
Transaction Size per Month	38,000,000
Data Entry Input Frequency	One tape every ten days
Programming Language for Update	FORTRAN, assembler
Estimated Program Size	Over 100 programs
Data Retrieval Method Used	FORTRAN, assembler
Master File Record Size	2404 for minute file
Retention Period for Master Files	Permanent
Average Number of Transactions per Update	12,500,000
Data Base Management System Used	None
SYSTEM COSTS	
Fiscal 75 Computer Costs	$117,000
Fiscal 76 Computer Costs	$249,762
Fiscal 74-75 Contractor Costs	Not available to Index
Fiscal 76 Contractor Costs	Not available to Index
Current Manpower Support	6 EPA 6-8 Contractors

REGIONAL AIR POLLUTION STUDY (RAPS)

SYSTEM MANAGEMENT	
System Documentation	Yes
Documentation Update Procedures	Yes
Contractors for System Operation	Yes
Contractors for System Maintenance	Yes
User Training Manuals	Yes

COMMUNITY HEALTH AIR MONITORING PROGRAM (CHAMP)

SYSTEM DESCRIPTION	
User	Health Effects Research Laboratory,ORD
Computer	UNIVAC
Facility	NCC

Although data gathering associated with CHESS has ended, data are are being analyzed by researchers for report publication. Fiscal 1975 computer costs were $231,000. The CHAMP program is in the data gathering phase. Currently, a feasibility study is under way to evaluate system designs for managing air quality data associated with CHAMP. However, cost data associated with the feasibility study and projected operating costs were not available to Index. An interim data system is operational for the capture of air quality data associated with CHAMP. The fiscal 1976 computer costs for this system and the continuing CHESS analysis are approximately $350,000. Other descriptive system and cost related information were not available for CHESS and CHAMP.

APPENDIX B - COST ESTIMATES

APPENDIX B - COST ESTIMATES

This Appendix presents the estimation rules and assumptions employed for evaluating alternative system actions and for projecting ADP expenditures resultinq from proposed changes. Section 1 describes the estimation procedures as applied to the evaluation of system design alternatives for managing air quality and emissions data. Section 2 describes the consulting and system development costs which result from proposed changes.

1. COST ESTIMATES FOR ALTERNATIVE SYSTEM DESIGNS ASSOCIATED WITH THE MANAGEMENT OF AIR QUALITY AND EMISSIONS DATA

Cost estimates utilized for evaluating alternative system design philosophies were obtained from four sources, as follows:

- Annualized data system operating costs from EPA TSSMS reports
- EPA system managers' ADP cost estimates
- Development and operational costs of similar EPA data systems adjusted for differences in data volumes and scope
- Cost estimates utilized in earlier studies performed by Index and other contractors, where appropriate

Cost estimates presented in this section do not include EPA personnel costs.

(1) Air Quality Data System Computer Operating Costs

The following cost estimates and assumptions were used to approximate air quality computer operating costs:

- Current SAROAD modifications have been estimated by the current systems contractor to reduce annual data retrievals and maintenance ADP costs by at least 50% to approximately $190,000 and $160,000, respectively
- Index estimates state computer billing rates to average $300 per hour
- EPA system managers have projected that 35 states will utilize either AQDHS-II or a locally-developed ambient air data system
- A typical state-operated system requires 30 hours of computer time to operate annually
- An EPA-operated data system for states that cannot support their own system requires 9 hours of computer time to operate annually
- The aggregate operating costs to EPA associated with the states that cannot support their own data system would be approximately $40,000
- State coordination and data handling costs associated with providing ambient monitoring data to EPA, quarterly, have been estimated by Index to be $50,000
- The operation and reporting associated with a summary or violations information system are expected to cost $35,000. This system would be similar in complexity to the Formal Reporting System
- Data communication costs for a utility approach would constitute approximately 20% of the annual maintenance and reporting costs

The following cost parameters were defined in order to estimate computer operating costs associated with the alternative data system designs:

S = aggregate state maintenance and retrieval costs = 30 hours/year x 35 states x \$300/hr = \$315,000

E = EPA retrieval costs = \$190,000

X = EPA maintenance costs for its data system = \$160,000

Z = EPA operating costs for states that cannot support their own data system = 9 hours/year x 15 states x \$300/hr = \$40,500

M = operating and reporting costs associated with management summary violations data = \$35,000

The computer operating costs associated with the current, hybrid and utility design approach have been estimated by Index to be represented by three equations, as follows:

Current	=	S + 1.5E + X	= \$795,000
Hybrid	=	S + 1.1E + 0.1X + Z + M	= \$615,000
Utility	=	1.5S + E + Z	= \$705,000

These equations are adjusted to account for the relative contribution of the previously mentioned parameters to the aggregated operating costs for each alternative.

The operating costs associated with the current approach include both EPA and state maintenance and retrieval costs (S, X, E). Because of completeness and currency deficiencies associated with the current EPA-maintained data base, data retrievals are supplied at present by states to EPA. With a hybrid approach data system, the operating costs to states would be similar to the current approach. On the other hand, EPA reporting and maintenance costs (E, X) would be reduced substantially. However, EPA would have to maintain air quality data (Z) for those states that cannot support their own systems. States would continue to supply data reports to EPA, as needed. Furthermore, with a hybrid design EPA would maintain a management reporting system (M) of summary and violations data. With a utility approach operating costs (S) to states would increase over the current and hybrid approaches because of data communication costs and additional costs associated with maintaining a central utility data system for a diverse group of users.

(2) Air Quality Data System Development and Installation Costs

The following cost estimates and assumptions were used to approximate air quality data system development and installation costs:

- Installation costs for AQDHS-II are approximately $17,000 per state

- An additional 10-15 states are expected to install AQDHS-II within the next five years
- Modifications to AQDHS-II necessary for the current and hybrid design approaches are estimated to total $100,000 - $125,000
- Based on the scope of work necessary to develop a utility-oriented air quality data system, development costs would range between $500,000 - $800,000
- The installation and user training costs associated with implementing a utility-oriented data system for 35 states would be approximately $175,000
- The development of a management reporting system for violations information would be similar in work scope to the Federal Water Supply System or Federal Reporting System and would cost approximately $50,000

The current approach would involve additional AQDHS-II installations and modifications and would cost approximately $330,000. The development and installation costs necessary for a hybrid approach would be similar to those costs incurred for the current design, but would involve the development of a management reporting system for violations and summary information. These costs would total approximately $380,000. System development, installation and training costs necessary for a utility design would cost between $600,000 and $1,000,000.

(3) Emissions Data System Computer Operating Costs

The following cost estimates and assumptions were used to approximate emissions computer operating costs:

- EPA system managers have projected that 35 states would utilize either EIS/PR or a locally-developed emissions data system
- A typical state-operated system requires 15 hours of computer time to operate annually
- Index estimates state computer billing rates to average $300 per hour
- Maintenance costs for the EPA maintained data system NEDS will remain stable at $30,000
- EPA retrieval costs from NEDS would increase by 50% to $150,000
- State coordination and data handling costs associated with providing machine-readable data to EPA semi-annually are estimated to be $25,000

The cost estimates and assumptions were utilized to estimate the computer operating costs for the current and utility data system design. The various cost parameters for each approach are illustrated below in Table 1.

Table 1

	Current	Utility
State maintenance and retrievals	$160,000	$240,000
EPA maintenance	$ 30,000	-
State retrievals for EPA	$ 50,000	-
EPA retrievals	$150,000	$150,000
Semi-annual reporting costs for states to EPA	$ 25,000	-
TOTAL	$415,000	$390,000

The computer operating costs for the current and utility approaches do not differ substantially considering the imprecision of the estimation parameters. The elimination of EPA maintenance, state reporting and retrieval costs would be offset approximately by an increase of maintenance and data communication costs with a utility design.

(4) Emissions Data System Development and Installation Costs

The following cost estimates and assumptions were used to approximate emissions data system development and installation costs:

- Installation costs for EIS/PR are approximately $17,000 per state
- An additional 10-15 states are expected to install EIS/PR over the next five years
- Modifications to EIS/PR necessary for the current design approach are estimated to total $175,000
- Based on the scope of work necessary to develop a utility-oriented emissions data system, costs would range between $500,000 - $800,000
- The installation and user training costs associated with implementing a utility-oriented data system for 35 states would be approximately $175,000

The current approach involves additional EIS/PR installations and modifications and would cost approximately $380,000. On the other hand, system development, installation and training costs necessary for a utility design would cost between $600,000 and $1,000,000.

2. CONSULTING AND SYSTEM DEVELOPMENT COSTS WHICH RESULT FROM PROPOSED CHANGES

Cost estimates utilized for projecting consulting and system development efforts which result from proposed changes were based on four factors, as follows:

- EPA system managers' ADP cost estimates
- Scope and magnitude of the consulting or system development effort
- Number of people and elapsed time appropriate for each project
- Travel time and associated expenses

For estimation purposes, we have assumed the following average manmonth rates:

Contractor-consulting project	$6000
Contractor-system development	$4500

Exhibit B-1, on the following page, shows the cost derivation of each of the proposed consulting and system development efforts.

Exhibit B-1
U.S. Environmental Protection Agency
Costs Associated With Proposed Changes

	Estimated Cost	Elapsed Time	Basis for Cost
Procedure and Requirements definition for distributed approach	$25,000	3 months	3.5 mm × $6,000 = $21,000 expenses = $ 4,000 $25,000
Management summary and violations data system development	$50,000	6 months	2 mm × $6,000 = $12,000 7 mm × $4,500 = $31,500 expenses = $ 6,500 $50,000
EDS system audit	$15,000	3 months	2 mm × $6,000 = $12,000 expenses = $ 3,000 15,000
Regional NEDS edit and data validation development	$10,000	2 months	2 mm × $4,500 = $ 9,000 expenses = $ 1,000 10,000
CDHS enhancements	$250,000	18 months	EPA system managers' estimates
NCC Technical Review	$50,000	4 months	7 mm x $6,000 = $42,000 expenses = $ 8,000 $50,000

APPENDIX C - QUESTIONNAIRE RESPONSE SUMMARY

APPENDIX C - QUESTIONNAIRE RESPONSE SUMMARY

The information contained in this appendix was obtained from a questionnaire prepared by the Standing Air Monitoring Work Group and the Management Information and Data Systems Division. This appendix summarizes eight categories of information reported on the questionnaire, as follows:

- Primary uses of air data
- Types of air data required
- Problems with air data, information flows and systems
- Scope of air data requirements
- Computer systems used
- Preferred system design philosophy
- Summary of respondants
- Approximate Annual ADP-related costs

Questionnaires were completed by various EPA Headquarters, Research and Regional offices. The findings from the questionnaires provided additional background information on the scope of air program data needs and various aspects of the results were incorporated in the data needs and analysis chapters of this report.

The responses have been summarized by five respondent groups, as follows:

- Office of Air Quality and Planning Standards (OAQPS)
- Office of Planning and Evaluation (OPE)
- Division of Stationary Source Enforcement (DSSE)
- Office of Research and Development (ORD)
- Regional Offices

Apparent contradictory statements may appear within a group as a result of the grouping of various responses. Also, the responses have been separated into the eight information categories listed above.

SUMMARY OF AIR MONITORING AND SOURCE DATA NEEDS QUESTIONNAIRE

	OFFICE OF AIR QUALITY AND PLANNING STANDARDS	OFFICE OF PLANNING AND EVALUATION	DIVISION OF STATIONARY SOURCE ENFORCEMENT	OFFICE OF RESEARCH AND DEVELOPMENT	REGIONS
Primary Uses of Air Data	.Determine air quality status (a) .Prepare annual trend reports (a,e) .Develop and revise emission factors (e, o) .Determine environmental impact of NSPS, NESHAPS, FMVES (a,e,o) .Evaluate need to revise standards and to control new pollutants (a,e,o) .Perform special studies (a,e,o) .Develop and test control strategies (a,e,o) .Determine national manpower and resource needs	.Determine trends in population exposure, actual and potential emissions (a,e) .Determine impact of new sources on ambient air quality (a,e,o) .Validate air quality models (a,e) .Establish daily ambient air pollution indicator (a)	.Plan and order future enforcement activities (a,e,o) .Determine compliance status (o) .Measure progress in emissions reduction (e,o)	.Develop and evaluate air quality and source models (a,e,o) .Research and pollutant assessment studies (a,o) .Identification of new pollutants (a,o)	.Revise SIP's (a,e,o) .New source review (a,e,o) .Enforcement actions (a,e,o) .Evaluate control strategies (a,e,o) .Modeling applications (a,e,o) .Air quality maintenance planning (a,e,o) .TCP development (a,e,o) .Non-significant deterioration (a,e,o) .Determine source compliance (e,o)

a = Air Quality Data
e = Emissions Data
o = Other Data

	OFFICE OF AIR QUALITY AND PLANNING STANDARDS	OFFICE OF PLANNING AND EVALUATION	DIVISION OF STATIONARY SOURCE ENFORCEMENT	OFFICE OR RESEARCH AND DEVELOPMENT	REGIONS
Air Data Needs	.Raw and summary air quality data (criteria and non-criteria) .Timely and complete air quality data (representing 90 days or more) .Area and point source emissions for urban areas .Industrial process data .Ambient site information .Short-term emissions data .Modal and diurnal data on source operations .Land use based emissions .Historical emissions	.Ambient and source data to allow isopleths to be drawn .Timely, complete and historical emissions data .Background and source data for criteria and non-criteria pollutants .Inclusion of new sources, planned technologies and construction schedules .Meteorological data	.Complete ambient data no more than 6 months old .Annual emisssions data .Compliance status	.Raw data for criteria pollutants, metals and toxic substances .Power plant emissions data .Meteorological data .Quality assurance information ."CAMP" type data for more non-urban stations .Sulfates and acid rain data .Particulate characterization	.Raw and summary air quality data .Meteorological data .Ambient data for non-criteria pollutants .Compliance data .Fuels data .Automobile emission data .Quality assurance information .Point and area source emissions inventories .Accuracy ($\pm$ 10%)
Scope of Air Data	National (some local data needs for modeling applications) Historical (5 years)	National (some local data needs for modeling applications) Historical (5 years)	National summary Historical data not necessary	National (some local data needs) Historical (5-20 years)	Regional and local Historical (5 years)

	OFFICE OF AIR QUALITY AND PLANNING STANDARDS	OFFICE OF PLANNING AND EVALUATION	DIVISION OF STATIONARY SOURCE ENFORCEMENT	OFFICE OF RESEARCH AND DEVELOPMENT	REGIONS
Problems Associated With Data, Information Flows and Systems	.Insufficient data accuracy, completeness and currency .Non-criteria data from more areas are needed .Limited software capabilities .Unstable hardware .Slow turn-around time .More complete data from fewer sites needed .Insufficient quality control by data collectors	.Less data needed per site but covering additional sites .Incomplete, out-of-date and poor quality data .Methods not standardized .Siting not always appropriate .Meteorological data not localized enough .Systems not designed to store special study data for future use .Limited state access to EPA data bases .Historical emissions not maintained .Ambient data presented on site-by-site basis are inadequate for interpretation	.Ambient data too old .Unnecessary steps in information flow process .Outdated and incomplete emissions data .Unstable hardware	.Quality of data unknown .Access to raw data difficult .Inflexible software and time delays in data flow .Limited mixing depth data available .More complete data needed for fewer locations .Unfamiliarity with UNIVAC software and operations	.Poor quality and outdated data .Insufficient meteorological data .Siting and calibration problems .Operational costs for CDHS not accurately projected .Slow turn-around time .No plot capabilities .No polygon retrievals .Greater fliexibility in retrieval programs needed .Training programs and technical assistance needed .Awkward and time-consuming data flows .ADP suballowance is not sufficient to allow increased NCC usage without reprogramming actions .Qualification of data by quality indicators needed

	OFFICE OF AIR QUALITY AND PLANNING STANDARDS	OFFICE OF PLANNING AND EVALUATION	DIVISION OF STATIONARY SOURCE ENFORCEMENT	OFFICE OF RESEARCH AND DEVELOPMENT	REGIONS
Preferred System Design Philosophy	.Centralized data bank with standard data formats, collection and quality assurance procedures	.Less lag time, more believable data, with Regional/state/local access (at cost)	.Highly decentralized system .Regional responsibility for operation of data system .Centralized processing at regional level with headquarters access to regional data bases	.Data collection, reduction, storage and retrieval automatic with little human interface .Hybrid system with in-depth data for Regions on all monitoring stations within their geography and smaller network of reporting stations with national representation for headquarters use .Hybrid system with decentralized data reporting and analysis	.There can be no one optimal system. Systems are designed for specific purposes .State access and update control of EPA data systems .Centralized data base and control .Decentralization of all systems operations where economically feasible .Partial interfaces among all systems .Data management program at each level of system usage .Hybrid system with limited trend data from a small number of stations being centralized and more extensive data from states handled by a decentralized regional system .Direct edit capability at NCC for states

	OFFICE OF AIR QUALITY AND PLANNING STANDARDS	OFFICE OF PLANNING AND EVALUATION	DIVISION OF STATIONARY SOURCE ENFORCEMENT	OFFICE OF RESEARCH AND DEVELOPMENT	REGIONS
Computer Systems Used	SAROAD, NEDS, CDS, RAPS, EOS, REPS, CAASE, SOTDAT, HATREMS	SAROAD, NEDS, CDS	SAROAD, NEDS, CDS	SAROAD, NEDS, RAPS other	SAROAD, NEDS, CDS State Implementation of CDHS
Respondants	Monitoring and Data Analysis Division Emission Standards and Engineering Division Strategies and Air Standards Division Control Programs Development Division	Office of Planning and Evaluation	Division of Stationary Source Enforcement	.Environmental Monitoring and Support Labs -Corvallis -RTP (2) -Las Vegas .Office of Health and Ecological Effects .Monitoring Technology Division	All Regional Offices

ESTIMATED ANNUAL ADP COSTS *

Organization	System	Computer $ (000)	Personnel $ (000)	Contractor $ (000)	Total $(000)
Office of Air Quality Planning and Standards Monitoring and Data Analysis Div.	SAROAD NEDS Modeling Other Total	900 350 200 75 1525	250 60 160 45 515	250 210 300 200 960	1400 620 660 320 3000
Office of Planning and Evaluation	SAROAD, NEDS, CDS				60
Division of Stationary Source Enforcement	CDS	65	1 Man Year	35	
Office of Research and Development Las Vegas					47

*Estimated annual ADP cost represent those costs stated on the questionnaire and do not reflect necessarily the costs as projected by Index Systems.

Organization *	System	Computer $ (000)	Personnel $ (000)	Contractor $ (000)	Total $ (000)
Region II		14.7	13.5		
Region III		15.0	2 Man Years	3.0	
Region IV	CDS Only	10.0	1 Man Year	50.0	
Region V					75.0
Region VI		10.0	25.0		
Region VII	SAROAD NEDS CDS AQDM CDM	2.7 1.6 12.6 .7 .4	4.6 4.5 12.8 4.4 4.4		7.3 6.1 25.4 5.1 4.8
Region VIII	SAROAD CDS	8.9 1.3			
Region IX	CDS AEROS Other	7.0 3.0 6.0	18.0 9.0 5.0	6.0 6.0	

* Region I and X cost estimates not received by Index Systems.