Developing a Statewide Emission Inventory Using Geographic Information Systems (GIS)

Patrick Gaffney
California Air Resources Board
Planning and Technical Support Division
1001 I Street, Sacramento, CA 95814
pgaffney@arb.ca.gov

U.S. EPA Annual Emission Inventory Conference
Atlanta, Georgia
April 2002

ABSTRACT

Geographic information system (GIS) technology can provide significant enhancements in estimating and analyzing emission estimates of airborne pollutants. These capabilities will help to improve our ozone and particulate matter attainment plans. In addition, GIS is an important tool for evaluating neighborhood level community health air pollution impacts. In combination with the Internet, GIS allow us to more effectively display our results to the public and help them understand the types and sources of air pollution around them.

In this paper, we provide our rationale for deploying GIS-based emission inventory systems for the Planning and Technical Support Division of the California State Air Resources Board.
A planning framework is also included for making concrete steps towards incorporating GIS into air pollution emission inventories within a governmental agency. The paper outlines software and hardware needs to support approximately 20-40 GIS users, as well as providing suggestions for staff training, and ideas for developing a GIS data library, which is the backbone of a GIS system.

The application of GIS for air pollution analysis is still relatively new and it requires a significant ongoing commitment to implement. Ultimately, emission inventories that incorporate GIS will substantially improve our ability to develop effective plans to meet air quality standards and help understand the effects of air pollution at the local community level. This paper provides suggestions for developing a plan and foundation for incorporating GIS technology into our emission inventories.

WHAT IS GIS AND WHY USE GIS?

Simply put, a geographic information system is a means of electronically storing, analyzing, and displaying data that innately includes a spatial component; it includes actual location information (and often much more). GIS data systems easily store and manipulate spatial objects such as areas, polygons, boundaries, lines, and points. Each of these objects relate to real-world features such as census tracts, facility locations, roads, rivers, elevation, spatial demographic information, and political boundaries, all of which can be combined, interrelated, and analyzed using GIS tools.

Within the Planning and Technical Support Division (PTSD) of the California Air Resources Board (ARB), we have embarked on an effort to incorporate GIS technology into our emission inventory systems. Our goal is to make incremental improvements to our emission inventory and source data using GIS, while also evaluating the benefits and shortcomings of a GIS based approach.

Why Create a GIS Based Emission Inventory?
Some of the primary motivations for creating a GIS based emission inventory include:

1. **Spatial Resolution.** Much of what we do for emission inventories includes a spatial component such as facility locations, traffic volumes, or agricultural field activities. Unfortunately, this information is often not available in a way that makes it easy to use or display in a spatial format, such as on a map. Now, the technology and computing power are readily available to fully incorporate spatial and temporal data into our emission inventory efforts.

2. **Accuracy.** It is increasingly important to provide accurate, detailed, and consistent emissions information to support the air quality planning and atmospheric modeling requirements for state implementation plans, environmental justice needs, and localized air pollution concerns. GIS is a very effective way to store and manipulate spatial data, verify its accuracy, and provide policy makers with thorough information necessary for improving public health.

3. **Consistency**. As emission inventories become more refined and more detailed, they become more complex. The information used to generate inventories also becomes more complicated and more difficult to share. By using geographically based information, we can develop consistent and interrelated inventory data that can be shared, compared, and analyzed in ways that are now not possible. Using GIS also allows us to more effectively combine and compare emissions data with air quality data, meteorology, and modeling results.

4. **Accessible and Understandable.** We have made substantial progress in making emission inventory data more accessible to our many users. However, there are pressing needs to provide the data in interactive and graphical formats. For example, it is important to provide emission maps that allow people to visualize and perform analyses

of air emissions using simple web based tools. A common request is, "Show me what emission sources are within one mile of my house?" For dynamic, online interactive requests, this capability can only be reasonably provided using GIS technology.

Within California, most of the straightforward sources of air pollution have already been controlled. Therefore, to continue making effective progress in reducing air pollution for our attainment demonstration plans, we must have a better understanding of when and where pollution occurs, and the effects of emission reductions. GIS provides an important tool in answering these questions. In addition, California is committed to evaluating environmental justice and community level air pollution impacts in all of our programs. Through the use of GIS technology, we can store and analyze neighborhood level emissions data, compare it with other information, such as income or race, and display the results interactively on easy to use and interpret maps.

PLANNING FOR A GIS EMISSION INVENTORY

Within the Planning and Technical Support Division (PTSD) of the ARB, we are gradually developing a GIS based emission inventory using an incremental approach that requires relatively modest resources, uses existing staff, and requires limited contractor support. It is worth noting that the ARB does not have dedicated GIS staffing or a GIS department, so all GIS work is being performed and managed by air pollution scientists who have shown a desire and aptitude for adopting the technology. For many of us, using GIS applications and products are only a small part of our overall responsibilities. Although this approach has some challenges, such as maintaining focus and progress while working on other priorities, the approach also has several benefits.

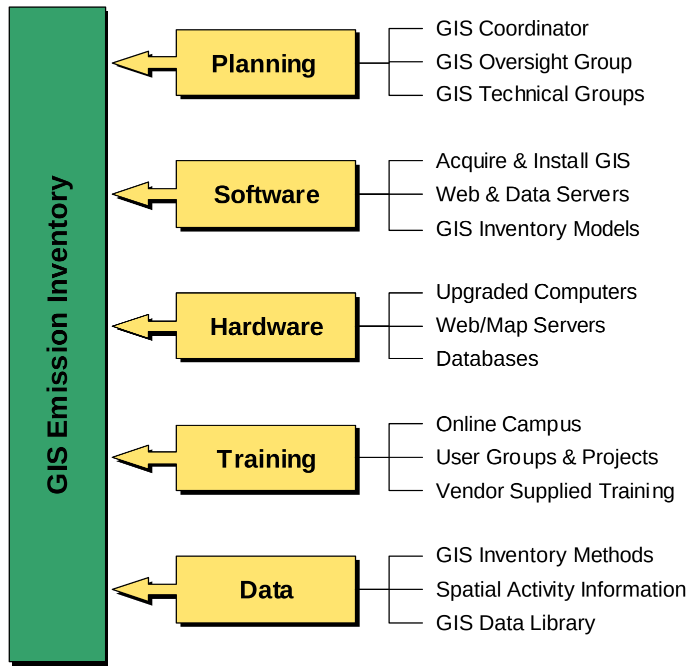
In planning our efforts, we also considered an alternative method of deploying GIS which included fully evaluating our GIS system needs and requirements, determining the final products, then preparing a detailed financial and labor plan for developing those products. Next, we would hire contractor support to develop the specified systems. We did not pursue this approach due to difficulties in determining what final products should ultimately be developed, the pathway to their development, and uncertainties in sustaining multi-year funding.

To date, we have focused on two phases of our GIS development. This paper primarily focuses on the Phase 1 activities, which are the steps used to build a foundation to provide the institutional support and tangible results needed to justify moving forward. Phase 2 of our work begins the process of more fully integrating GIS into our emission inventories, and includes more complex and ambitious projects requiring additional financial and staffing resources.

Phase 1 GIS Planning

Successfully deploying GIS software to a large group and making it useful requires several components. Software must be acquired, hardware upgrades may be necessary, staff training is required, data must be developed and stored, and an overall planning and coordination structure is needed. Figure 1 summarizes these components.

Figure 1. GIS Emission Inventory Components



Phase 1 initiated development of the basic structures necessary to develop GIS system for the ARB emission inventory group. This phase required modest financial resources and no new staffing. However, it required ongoing time commitments from existing staff. Most of the Phase 1 objectives were completed within six months. Phase 1 included the following components:

- Assign a GIS coordinator and develop a leadership structure
- Create technical teams for evaluating software licensing, hardware needs, training needs, and a data library
- Acquire the necessary GIS software and training
- Acquire, configure, and maintain the hardware/software systems necessary for providing map based GIS data via the internet
- Develop prototype systems and mockups showing the benefits and functionality of using GIS
- Provide demonstrations and obtain management approval for fully developing the most promising prototype systems
- Develop a preliminary spatial data library that can readily be shared across the division and within the ARB
- Host GIS coordination and technology sharing meetings
- Conclude Phase 1 with presentations to Executive Officers and other management; Prepare Phase 2 plan

Phase 1 provides tangible products, while also giving staff operational, hands-on experience using GIS. By increasing our baseline knowledge, we can clearly identify the capabilities and limitations of GIS as we move to the more extensive Phase 2 projects. A key objective of Phase 1 is to provide the ARB leadership concrete applications that demonstrate what GIS technology is, what it can do, and how it will be applied to achieving our air quality goals.

Tracking Progress and Moving Forward

In the past, GIS had a tendency to fall through the cracks as other priorities overwhelmed it. Therefore, one of our initial planning steps was to develop a clear progress tracking structure. Developing a clear plan and schedule has been essential in helping management and staff to clearly commit to the new workload demands necessary for developing GIS-based inventories and methods. To ensure our success, it was important to:

- Ensure management team buy-in and accountability for results
- Identify motivated team leaders and provide them with adequate resources
- Require regular meetings and status updates within groups and with management
- Assign staff tangible projects and milestones and allocate time for work
- Require staff GIS training as part of individual development plans
- Strive to ensure GIS progress while maintaining baseline non-GIS workloads
- Identify areas where contractor support can bolster internal efforts

To ensure that we meet our goals the table on the following page (Table 2) provides a schedule and checklist used to track progress in meeting our GIS inventory objectives. Referring back to the GIS inventory components diagram in Figure 1, these tasks are the projects necessary to complete the hardware, software, training, data, and planning requirements shown.

GIS Inventory Planning and Technical Groups

We also found it valuable to create a set of planning and technical groups. Each group is assigned a lead staff person accountable to a manager. Groups are responsible for the objectives listed in Table 1, and for meeting various objectives listed in the Table 2 schedule.

We also assigned small GIS related tasks to staff to create tangible GIS products that concretely improve the emission inventory. Initially called 'GIS-mini-projects,' they were developed to help educate the staff, as well as to provide clear demonstrations of the benefits of using GIS for emission estimates.

Table 1. GIS Planning and Technical Groups

Group	Responsibilities
GIS Inventory Planning Group	Overall Coordination Enforce Schedules Project Planning
User Support and Training	On-Line Training Users Group Technical Support License Management
Division Data Library	Standards & Organization Maintenance Geocoded Data
GIS Projects and Tools	Internal Projects Outside Research Prototype Development
GIS Spatial Database System	Feasibility Study Report Air District Outreach System Design
Emission Inventory Development	Mobile Source Emissions Point & Areawide Sources System Integration System Evaluation & Demo

The mini-projects included developing an initial GIS based inventory for gas stations, creating a demonstration inventory for emergency standby generators, preparing an audit of existing facility location data, and locating air districts and district rules using GIS. In addition to these projects, larger ongoing projects are in place to develop a GIS community health air pollution system, estimate biogenic emissions using GIS, compute ammonia emissions for fertilizers and soil, and develop road-link based emission estimates for on-road vehicles.

Table 2. GIS Emission Inventory Milestones for 2001-2002

Project Description	JUN 2001	JUL	AUG	SEP	OCT	NOV	DEC	JAN 2002	FEB	MAR APR	MAY JUN	JUL AUG
Set-Up and Planning												
Write Initial Scoping Plan		☑										
Acquire Software		☑										
Acquire Hardware & Upgrades			☑									
Acquire Internet Map Server				☑								
Establish Work Groups				☑								
Provide Training				☑						☐		
Create Initial Data Library					☑					☐		
Planning and Strategy Group Meetings				☑	☑	☑	☐	☑	☑	☐	☐	☐
Prototype Projects and Systems												
Community Health System							☑			☐		
Biomass Utilization Prototype									☑			
Air Quality Display System							☑					☐
Biogenics Emissions Model												☐
Fertilizer Ammonia Model											☐	
Prototype Emission Inventory Development (excerpts)												
Prepare Work Plan									☑			
Performance Reviews with Management										☐		☐
Update Emissions Using GIS											☐	
Evaluate Spatial Database Designs											☐	
System Integration & Consistency											☐	
System Evaluation & Demonstration												☐
Identify Future Projects												☐

☐ Project Milestone
☑ Milestone Completed

The remaining sections of this paper provide technical details regarding the software, hardware, training, and data needed for developing a state-level GIS based emission inventory system.

SOFTWARE

Software is the foundation for implementing a geographic information system. These are the programs that allow computers to store, retrieve, manipulate, and analyze the points, lines, polygons, rasters, and other components used to represent facilities, roads, counties, elevation profiles, and other real-world components.

Within the ARB, selecting a GIS software vendor was easy because the California Environmental Protection Agency, which includes the ARB, standardized on using ESRI™ (Environmental Systems Research Institute) GIS software. Although there are many other GIS software vendors, it was important for us to rely on the current 'industry standard' so we have a high level of internal consistency, as well as having consistency with other California state agencies.

Software Needs

In selecting the types and quantity of GIS software needed, we had two main considerations. The first consideration was to provide GIS access to a large number of people. We estimated that within a year, we could expect 20-40 people to have a consistent need to use GIS in their work. The second consideration was cost. Our initial configuration had a cost of nearly $60,000 (just for software). This was ultimately scaled down to about $35,000, which included discounts from upgrading existing software.

For hosting a large number of potential GIS users, a factor that allowed substantial cost reductions was the use of what are called 'concurrent' software licenses. Using concurrent licensing, anyone who wants to can install the software on his or her computer. However, when they try to use the software, it checks with a network license manager to see if there is a 'seat' available that has not been checked-out. Concurrent licenses cost about three times more than a single computer installation software package, but when there are many users who do not use GIS full-time, this is the most cost effective approach for configuring desktop GIS systems.

Purchasing GIS software can be significantly more complicated (and expensive) than typical software packages. For example, the core ESRI GIS software is available as three versions: ArcInfo provides the greatest capability at the greatest cost and is typically for higher end applications and analysis, ArcEdit is the mid-level product, and ArcView is the standard desktop software which meets the needs of most GIS users. In addition, extensions to the core software are available that provide additional functionality. For example, Spatial Analyst is used to work with gridded data, spatial statistical analyses are performed with Geostatistical Analyst, 3D Analyst is for working with three-dimensional information, and address matching is performed using ArcGIS Streetmap. Naturally, this functionality comes at an additional cost, which is typically as much as the core ArcView software.

As mentioned, the situation is further complicated because most of the software is available as either a single user license (one computer), or as a concurrent license (at additional cost) that allows the software to be 'checked-out' over a network by multiple users. ESRI GIS software also has the option of entering into annual maintenance agreements that provides software upgrades and technical support. In addition, existing users of ESRI software can upgrade to current software versions as well as to concurrent licenses at a reduced cost from purchasing a new products.

In ordering our software, we wanted to support a relatively large number of casual users (20-40) and several 'power-users'. Instead of ordering incrementally, we acquired enough licenses to fully support our GIS development for at least a year. This lead to a complex order, which is probably more than most people would need to get started.

In summary, we decided to completely convert over to the latest ESRI GIS software suite, which is named ArcGIS™ 8. Our current software configuration includes the following concurrent licenses: 10 ArcView licenses, 2 ArcInfo, 3 Spatial Analyst, 2 Geostatistical Analyst, and 2 ArcGIS Streetmap. We also ordered three sets of single user ArcView and some extensions to install on laptops or for use if all of the concurrent licenses are utilized and a high priority project needs access to the software.

HARDWARE

With software selected, the next step was to acquire hardware to support it. Current desktop GIS software can be effectively run on a reasonably well outfitted desktop computer. Within the ARB, most of the GIS users have computers with a 1-gigahertz processor, 128-256 megabytes of random access memory, a 20-40 gigabyte hard drive, Windows 2000, and attachment to a local area network. Fortunately, as we deployed our GIS software, the ARB was also performing agency-wide upgrades to our desktop computers, so additional funding was not needed for GIS desktop hardware.

In addition to desktop systems, we also purchased additional hardware and software to support serving GIS based maps interactively over the Internet. Because of the additional processing, data storage, and networking functionality, this system requires a server-grade computer that includes a dual processor, extensive storage space, and all of the necessary software to support serving interactive maps to the web. The cost for this Internet map system was about $25,000 for hardware and $25,000 for software and licensing. We are now in the process of acquiring additional hardware and software to develop and test a large-scale spatial database using either Oracle or DB2.

Unlike in the past, when dedicated, high-end GIS computers were required, computer hardware is now rarely a limiting factor in implementing desktop-level GIS systems. Relatively basic systems have the speed and capacity to handle a wide range of desktop oriented GIS activities.

TRAINING

A year ago, five people within the ARB knew how to use GIS software. Now there are at least 15 users and the number is growing. There are a couple of reasons for this. First, anyone who wants it can now have GIS software on his or her computer. Second, there is more demand for GIS related products, analysis, and maps, so staff has had to learn and use GIS in their work. Within the ARB's Planning and Technical Support Division, we are using several types of training, which are listed below.

Self-Study Tutorials and Virtual Campus

Provided with the purchase of ESRI GIS software is an extensive set of printed users guides. These guides include very good tutorials that introduce the software capabilities and are an excellent starting point for new users. In addition, self-paced online training is available from ESRI that provides comprehensive education. These courses are extremely cost effective ($100/class) and are well designed, but we are finding that many people are having difficulty consistently blocking-out the time needed to take the classes (20-24 hours) while maintaining their existing workloads.

In-House GIS Mini-Projects and User Groups

One very particularly effective training technique is requiring staff to develop and complete small, short-term (2-4 months) GIS projects to make tangible emission inventory improvements. This approach is based on our experience that it is difficult to learn GIS and retain that knowledge unless it is used for real-world projects. Some of the completed 'mini-projects' included investigating emissions from gas stations, inventorying diesel standby generators, a validation of facility location information, display of airport activity volumes, and analysis of link-based vehicle traffic.

At the completion of the original mini-projects, staff presented their projects and techniques. Not only was this educational, but it was also entertaining and helped to improve cohesiveness among the GIS users. We are now in the process of setting up user-based training in which staff members provide instruction, guidance, tricks, and tips on topics specifically selected to assist with issues of emission inventory development. Some planned topics include mapping latitude and longitude coordinates, using global positioning system (GPS) data, importing tabular data, labeling maps, buffering, gridding emissions data, and working with census data. Because of its in-house, subject specific focus, we expect that this training will be very helpful to the staff.

Other Training Options

There several other GIS training options including community colleges, colleges and universities, and vendor provided training. Each approach has benefits and weaknesses. Colleges can be inexpensive, but require a long-term commitment. Vendor training is typically very intensive (2-5 complete days), but costly ($400/day). Regardless of the selected training approach, we have repeatedly found that it is critical for people to have actual GIS projects to work on to supplement and reinforce the training. Without real-world experience and practice, even the best training quickly fades and becomes irrelevant.

GIS DATA LIBRARIES

The backbone for a GIS system is data. A GIS system without data is like a car without an engine or fuel. It may look nice, but you are not going to go anywhere. In order to be productive with GIS, it is important to develop consistent and accessible spatially enabled data that can easily be shared, compared, and analyzed.

Our first step in developing a GIS data library was establishing a GIS directory on one of our shared network hard drives. We have populated this directory with basic framework data that includes county boundaries, air basins, road networks, political boundaries, hill shaded relief maps, waterways, schools, hospitals, land use, vegetation, census data, and more. This information provides baseline information for those generating maps and analyses. An example of some framework data for California is located at: http://www.gis.ca.gov/data_index.epl

Another important step in establishing a data library is providing data standards. For example, we have established that all of our work will be done in the Teale Albers geographic projection. By specifying a projection up-front, we ensure the most complete and accurate data sharing possible. In addition, much of the GIS emission inventory information will ultimately be spatially gridded. To avoid the problems associated with different projects using grids with different specifications, we have established a unified one by one kilometer (1x1 km) grid for the entire state to be used for all future work.

We are now in the process of developing data sets beyond the framework data. This includes roadway link-based traffic volumes, health risk maps, detailed facility location and emissions information, gridded biogenic emissions, neighborhood emission sources, and spatial allocation of areawide sources of emissions such as agricultural tilling or prescribed fires. The results of this work provide consistent spatial emissions data as grids, point data, line data, or polygon data.

Using the power of GIS technology, any of these consistent spatial data can then be overlaid, added, compared, or analyzed against other information such as demographics or land use. By using GIS intelligently and developing an accessible and robust data library, we can reduce duplication of efforts, build incrementally on past efforts, and analyze data in ways that is not possible without the visual and spatial analysis capabilities available with GIS.

CONCLUSIONS

We began our GIS efforts in earnest about a year ago and are already seeing substantial payoffs. For example, a year ago we had only a few GIS capable staff. Now we have over fifteen staff members using GIS in their work to answer important questions about environmental justice, facility emissions, toxic risk, and attaining air quality standards.

Incorporating GIS technology into a large existing statewide emission inventory system is a substantial undertaking. New methods must be developed, new partnerships formed, new hardware and software acquired, new databases designed, and new information gathered. However, by using an incremental approach, the resources needed to provide the foundation for developing GIS emission inventory capabilities have been relatively modest.

In general terms, to host 20-40 GIS users, our software costs were approximately $35,000 (which included discounts for upgrades). It is expected that annual software maintenance costs will run around $15,000. Fortunately, our GIS rollout coincided with the rollout of new computers for all staff, so additional funding was not needed for desktop computer upgrades. Deploying the hardware and software needed to serve maps over the Internet cost approximately $50,000 ($25,000 for each component). Funding was also allocated to contracts for developing GIS Internet based mapping tools (approximately $150,000), which require specialized expertise not available within the ARB. For staff GIS training, costs could range from $3000-6000^{+}, depending on needs.

In terms of labor, it is difficult to estimate the person hours required in deploying GIS because GIS has become integrated onto some of our baseline workload and we do not have dedicated full-time GIS staff. Very roughly, over the past year, three managers allocated 5-10% of their time to GIS development, roughly 10-15 staff members have dedicated 5-15% of their time to GIS work, and 3-4 staff apply 80^{+} % of their time to GIS work.

With strong management support, we have now completed our Phase 1 infrastructure building efforts. We are now moving into Phase 2, which involves developing a spatially refined emission inventory for two small study regions within Southern California. The inventory will include high-resolution temporal and spatial data for mobile, point, and areawide sources of emissions. The goals of this phase are to begin developing, integrating, and updating the data and tools needed for preparing a statewide GIS based emission inventory, while also evaluating the strengths and weaknesses of a GIS based approach.

The ultimate goal of our GIS efforts is to improve air quality within California. Using GIS, we can meet this goal more cost effectively because will know more about when and where the emissions occur, and how they can be reduced to benefit the most people. Using GIS, we can improve air quality in those areas that are disproportionably affected by air pollution because we can analyze detailed information about neighborhood level emissions and toxic risk. Using GIS technologies, we are already seeing that we are able to provide more information, more clearly, for more people, at existing staffing levels.

Within several years, GIS will be just another tool we use like spreadsheets and databases, rather than the somewhat exotic entity it is now. Though our ongoing work, we will use GIS technology to combine and analyze emissions data with air quality data, meteorology, geography, demographics, and model results. The results of these efforts will provide a better understanding of air pollution within California, and cleaner air for those who live here.

DISCLAIMER

The opinions, findings, and conclusions expressed in this paper are those of the staff and not necessarily those of the California Air Resources Board. In addition, the mention of software products or vendors is for informational purposes only, and is not a recommendation or endorsement of the mentioned products.

KEY WORDS

GIS, emissions, spatial, inventory, planning