

Developing Internet-based GIS Applications to Manage and Disseminate Emissions Activity Data

Tami H. Funk

Dana L. Coe

Sonoma Technology, Inc.

1360 Redwood Way, Suite C

Petaluma, CA 94954

tami@sonomatech.com, dana@sonomatech.com

ABSTRACT

Geographic information system (GIS) technology has become an integral component of environmental management solutions. A GIS combines geographic feature information with spatially resolved data enabling display, analysis, and modeling of environmental data. Internet-based GIS applications provide a resource to display and disseminate air quality and emissions data to broad audiences. During the California Regional PM₁₀/PM_{2.5} Air Quality Study (CRPAQS), several types of databases were created to characterize the local vicinities of two dozen selected air quality monitoring sites. Field observers were dispatched to perform detailed GIS-based ground-truths of the land use/land cover and potential emissions sources within 2 kilometers of these 24 sites. In addition, field technicians maintained detailed logs of observed conditions in the vicinities of these sites throughout the study period. At one key site, digital cameras were deployed to collect a nearly continuous, 360-degree visual record of the surroundings. All information will be shared with numerous research teams, who are now responsible for analyses and interpretations of the air quality and meteorological data that were collected during CRPAQS. ArcIMS and ColdFusion were used to create a GIS-based Internet application, through which all the research teams can browse and query GIS data, logs, and an on-line archive of digital photographs.

INTRODUCTION

The California Regional PM₁₀/PM_{2.5} Air Quality Study (CRPAQS) is a multi-year program involving meteorological and air quality monitoring, emission inventory development, data analysis, and air quality simulation modeling. The field phase of CRPAQS included "annual" monitoring throughout the San Joaquin Valley and surrounding regions from December 1999 through January 2001, and "intensive" measurements during fall and winter-like conditions when PM₁₀ and PM_{2.5} concentrations are highest. The emission inventory development phase of CRPAQS included collection of emissions activity data during the 1999-2001 period to improve the California Air Resources Board's (ARB) emissions activity databases.

In support of emissions activity database development, Sonoma Technology, Inc. (STI) gathered and processed information from public records, commercial databases, and government sources to characterize activity patterns throughout the CRPAQS domain (Figure 1). Additionally, STI conducted visual and photographic observations of land features and emissions events in the near vicinities of selected monitoring sites. From these efforts, STI compiled ArcInfo and Microsoft Access-based

spatially resolved activity databases that are compatible with Geographic Information System (GIS) software.

These work products will be used to perform a variety of tasks during the data analysis phase of the CRPAQS. The domain-wide activity database includes detailed spatial and temporal data for a variety of emissions sources (e.g., gasoline service stations, bakeries, etc.). STI used the spatial data to produce gridded fields of area source surrogates and temporal profiles that may be used to spatially or temporally allocate area source emissions and generate input files for regional air quality models.¹ The site-specific/event-specific database may be used to enhance interpretation of air quality data from corresponding sites (especially for analyses of variability and/or unusually high ambient concentrations); evaluate site characteristics (e.g., urban vs. rural); characterize significant local emissions impacts on these sites; corroborate temporal activity profiles; and validate regional scale databases of emissions activities.

The emissions activity information collected will be shared with numerous research teams, who are responsible for analyses and interpretations of the air quality and meteorological data that were collected during CRPAQS. The Environmental Systems Research Institute's (ESRI) GIS Internet Mapping Service software, ArcIMS, was used in conjunction with the web programming language ColdFusion to create an Internet application, through which all the research teams can browse and query GIS data, logs, and an on-line archive of digital photographs. The Internet application is called the CRPAQS On-line Atlas with Regional and Site-specific Events (COARSE). The COARSE application is a web site with interactive mapping tools and options that allow users to explore and download information from the CRPAQS emissions activity databases.

TECHNICAL APPROACH

The COARSE Internet application enables users to query and display emissions activity data for 24 CRPAQS monitoring sites. The application interface includes mapping, navigation, and data query tools as well as access to data archives (i.e., on-line movie archives). Prior to developing the COARSE application, the spatial emissions activity data collected during the CRPAQS were placed into ArcGIS format, processed, and quality assured. The tabular emissions activity data include attribute information about the features in the spatial dataset and are contained in Microsoft Access databases. Subsequent to data processing and quality assurance, ArcIMS was used to compile the spatial map layers used by the COARSE application. The COARSE system design was established and the interface was developed to include the desired functionality. The remainder of this section provides a discussion of COARSE application development.

CRPAQS Emissions Activity Data Collection

The CRPAQS emissions activity data collection phase involved gathering and processing (1) short-duration emission events data that may impact CRPAQS air quality measurements and (2) highly resolved activity data in the immediate surroundings of selected CRPAQS monitoring sites. These data were gathered by observers stationed in the field, automated digital camera equipment, and publicly available records and/or news reports.

Land-use patterns were characterized in the 2-km x 2-km areas surrounding 24 CRPAQS monitoring sites. (The 2-km x 2-km area corresponds to the resolution of the CRPAQS air quality

modeling grid.) Field technicians utilized a customized, GIS-based data entry system to record the positions and spatial extents of land uses (e.g., agricultural, industrial), the positions of individual emissions sources (e.g., restaurants, gasoline service stations, unpaved parking lots), and further descriptive details (e.g., business hours of operation, stage of farming operations, etc.). The data entry system, the Local/Episodic Activity Data System (LEADS), comprises a Visual Basic 7.0 user interface that interacts with and/or populates Microsoft Access and MapInfo databases and a global positioning system (GPS). Using LEADS, field technicians defined and located potential emissions sources as points or areas on an electronic map display (e.g., restaurants, agricultural fields, residential areas with chimneys). A GPS was used to independently verify coordinates.

Field technicians performed data entry in the field using a laptop computer. Each of the 24 monitoring sites was visited twice during CRPAQS—once in the spring and again in the fall or winter. During the initial site visits (spring 2000), the field technician spent approximately 8 to 12 hours per site to fully characterize the land-use patterns and points within the 2-km x 2-km square area surrounding each site. During follow-up visits (fall/winter 2000/2001), the field technician spent 4 to 8 hours per site to revise and update the land-use patterns.

The 24 monitoring sites selected for ground-truth surveys are listed in Table 1 and illustrated in Figure 1. The sites listed in Table 1 were selected for ground-truthing activities because (1) they were likely to be impacted by the source types of principal interest (as judged by their spatial locations); (2) a variety of air quality and meteorological parameters were being measured at each site (e.g., chemical speciation); and/or (3) the sampling time scale was short for these sites; therefore, evidence of local sources was more likely to be apparent in the data. In addition, sites that demonstrated a correlation between measured concentrations and wind direction, which may indicate the influence of local emissions sources, were favored. Site maintenance personnel were trained to record changes in the surroundings whenever routine site maintenance visits were scheduled from spring 2000 through winter 2000/2001. Provided with note sheets and LEADS-generated area maps, site maintenance personnel traversed a circuit around each site and noted changes in the surroundings. The LEADS database was updated periodically based on the information from site maintenance personnel.

An elevated, four-way, time-lapse camera system was deployed at the Angiola monitoring site. The purpose of the camera system was to capture a continuous, 360° record of short-duration activities in the vicinity of the Angiola site. The Angiola Tower was selected as the ideal site for extra observational monitoring because, in addition to the qualities that it shared with the sites listed in Table 1, it had a clear view in all directions, a trailer to house and operate computer and electronic equipment, frequent attention from field staff members, and adequate site security. The camera system was deployed in late spring 2000 and operated continuously throughout fall and winter of 2000/2001. The captured images were indexed, archived, reviewed, and interpreted for LEADS data entry. All digital images were indexed by date and time with a standard file naming convention. Approximately one-third of the full image archive was reviewed and interpreted to record any observed emissions activities or changes in the surroundings with LEADS, and link the LEADS database to a corresponding image or series of images. Thus, end-users of the LEADS database may automatically retrieve and display images of interest. Additionally, a selection of images was reviewed in order to establish a baseline for comparison between days when local impacts were potentially present and days when local impacts were unlikely.

On a periodic cycle and during PM episodes or in response to unusual readings at specific monitoring sites, public records and news reports were monitored to identify short-duration events that

may have impacted air quality measurements. An electronic log of unusual events was maintained to augment the emissions activity databases.

Data Compilation, Formatting and Quality Assurance

When the ground-truth databases were complete, the spatial dataset was converted to ESRI's ArcGIS format. Spatial and temporal summaries of the emissions activity and ground-truth data were prepared in tabular and graphic form. Tabular and graphic summaries of emissions events and source locations were prepared in Microsoft Access and ArcView Shapefile formats. Data files were organized by county, sub-region, monitoring site, time of day, and source category. Graphic displays of the spatial allocation of emissions activities surrounding each of the 24 monitoring sites were generated using ArcView. The quality assurance (QA) process involved human evaluation aided by computer graphical and statistical tools. The QA review included (1) an evaluation of the emissions activity data as a whole and broken down by individual source categories to confirm the reasonableness of emissions activity and (2) an assessment of data formats and values to ensure the integrity of data files.

After the data quality assurance process, the Atlas of CRPAQS Ground Truth Surveys was prepared.² The atlas is an organized collection of graphics and tabular data containing all emissions activity information collected during the CRPAQS ground-truthing period. Figure 2 and Table 2 are excerpts from the Atlas of CRPAQS Ground Truth Surveys and provide examples of the type of activity information collected at each of the 24 monitoring sites.

Design and Development of the COARSE Application

Because the emissions activity data collected during the CRPAQS will need to be shared with numerous research teams who will be conducting analyses of the air quality data that were collected during CRPAQS, a custom application was designed that will enable a broad group of users to easily access this information via the Internet. The emissions activity dataset contains a great deal of spatial information including site maps, emissions source locations, and land use data as well as an extensive library of tabular data. One way of creating an application through which users can interactively navigate map layers and retrieve data is to use ESRI's ArcIMS software. ArcIMS is an Internet mapping software package that provides a framework for accessing GIS services and data on a Local Area Network (LAN) and on the Internet. Figure 3 illustrates the ArcIMS architecture and communication pathways among the various software components. ArcIMS allows multiple users to easily access, query, download, and disseminate geographic information through a standard web browser and can be customized in both appearance and functionality to satisfy user preferences and needs.³

The COARSE application is designed to run automatically with some modest manual intervention by the system administrator. Figure 4 shows a schematic of the COARSE application and data flow. As shown in Figure 3, the COARSE application, all spatial data files, and all Access database files used by the application reside on one web server computer. The application interface and functionality was developed and customized using the web programming language, ColdFusion. The map layers used by the application are static ESRI Shapefiles with pre-defined renderings established within ArcIMS. Queries that access the emissions activity databases are programmed in ColdFusion.

Design Issues

Prior to developing ArcIMS applications, several hardware and software specifications must be addressed based on the user audience, computing resources of the average user, and the desired functionality of the application. The ArcIMS viewers determine the functionality and graphical look of the site. The ArcIMS connectors provide a communication pathway between a web server and the ArcIMS Application Server. ArcIMS provides connectors for Java Servlet, ColdFusion, and ActiveX. The choice of connector affects the type of viewers that can be developed. All three support the built-in HTML viewer, which uses HTML/DHTML, Java, and JavaScript. Custom viewers can also be written using ColdFusion and ActiveX languages. Both ColdFusion and Microsoft's Active Server Pages (ASP, using ActiveX) are server-side processes, running on the web server. These interpret web page templates (*.cfm, *.asp) and return dynamic content. When used with ArcIMS, these programs can add functionality and provide a "thinner" client, with more processing being performed by the web server. ArcIMS applications are Internet-browser specific; therefore, the browser and browser-version compatibility should be established early in the design process.

The COARSE application was designed to function on a web server running Microsoft Windows 2000 within a firewall. The application was built within a customized ColdFusion-based viewer and the ColdFusion connector. The ColdFusion was selected for use in this project because it is a powerful yet versatile tool. The advantages of using ColdFusion are rapid deployment (translating into more features for development resources), performance, stability, and cross-platform support.

Features and Functions of the COARSE Application

The user interface for the COARSE application is shown in Figure 5. The interface provides users with navigation tools, data query and download options, and mapping tools. The COARSE application was designed to provide an interface that remains consistent as users navigate the site and perform queries; the data options and queries may change, but the same general page layout is preserved. COARSE is still in the developmental stage and it is likely that modifications will be made to the interface prior to application completion. The remainder of this section provides a general discussion of each of the interface components.

Navigation Tools

The navigation toolset, located in the upper left-hand section of the interface, allows users to navigate the map and zoom into geographic regions or monitoring sites of interest. Regions of interest can be defined by using the dropdown selection menu for sites and counties or by using the tools (icons) available in the navigation toolbar. When a site or a county is selected from the dropdown list, the map window automatically zooms to the selected region. The tools included in the navigation toolbar function by selecting an icon and manually clicking on or selecting features within the map window.

Activity Reports

The "Activity Reports" menu, located in the middle left section of the interface, provides two options for viewing and downloading emissions activity reports for the selected geographic region (region visible in the map window). The "Short Report" button produces a downloadable text file containing all the emissions facilities and land-use types visible in the map window including information about each source and land-use type. The "Long Report" button results in a downloadable

text file containing the same information as in the short report, with additional available emissions activity information (e.g., facility hours of operation).

The “Site Logs” button retrieves site logs for those sites visible in the map window. The site logs are diaries containing notes and information about emissions events and activities observed by field technicians during the CRPAQS period. The “Emissions Events” button retrieves a map layer showing the locations of emissions events that occurred during the CRPAQS period. Information about these emissions events is based on archived news articles and public records. The “Angiola Photos” button allows users to view digital photos collected at the Angiola monitoring site during the CRPAQS period. These photos can be viewed for selected dates and time periods of interest.

Map Layers

The “Map Layers” list, located in the upper right section of the interface, allows users to toggle map layers on and off. Selected map layers contain feature attribute information such as feature names and identifiers. This information can be accessed using the “Active Layer” selection window located in the bottom right section of the interface. When a map layer is selected from the active layer dropdown list, the selected layer becomes active and attribute information about map features can be obtained by clicking on the feature of interest in the map window.

Related Activities

There are numerous large scale GIS projects underway with the objectives of standardizing, managing, and distributing geographic data effectively. One example is the Open GIS Consortium, Inc. (OGC), an industry consortium aimed at growing interoperability for technologies involving spatial information and location. The following is an excerpt from the OGS vision statement:

*“Approximately 80% of business and government information has some reference to location, but until recently the power of geographic or spatial information and location has been underutilized as a vital resource for improving economic productivity, decision-making, and delivery of services. We are an increasingly distributed and mobile society. Our technologies, services, and information resources must be able to leverage location (i.e., my geographic position right now) and the spatial information that helps us visualize and analyze situations geographically. Products and services that conform to OGC open interface specifications enable users to freely exchange and apply spatial information, applications and services across networks, different platforms and products.”*⁴

There appears to be a growing interest in the use of GIS-based tools to manage and display spatial data. STI recently completed work on an Environmental Protection Agency (EPA)-funded proof of concept project, *AIRNow Interactive*, to develop an ArcIMS-based application to demonstrate how Internet-based GIS technology can be used to enhance the AIRNow system. STI is currently under contract to CARB to develop a new ArcIMS-based web interface to the Air Quality and Meteorological Information System (AQMIS), which will display and disseminate real-time air quality and meteorological data contained in the AQMIS database.

CONCLUSIONS

Initially, GIS was used as a tool for integrating, analyzing, and visualizing geographic information. GIS later emerged as a powerful technology for coordinating and integrating the work of an organization around a common and shared geographic database. GIS is now evolving into an architecture allowing communities of organizations to connect and share their geographic information and applications as network and Internet-based services to one another.⁵

As GIS software becomes increasingly user-friendly and the availability of geographic data continues to grow, custom GIS applications will become incredibly useful for managing, displaying, analyzing, and sharing environmental data. Custom GIS applications provide a cost-effective way to create custom database interfaces, analysis tools, and data dissemination tools for a broad audience of users. ArcIMS-based applications are scaleable and can incorporate many different types of data from different sources (databases). Combining ArcIMS with the programming flexibility of ColdFusion (or other compatible web development languages) allows for nearly limitless application customization.

REFERENCES

1. Funk, T. H.; Stiefer, P.S.; Chinkin, L.R. Development of Gridded Spatial Allocation Factors for the State of California; Technical memorandum, Prepared for the California Air Resources Board, July 2001.
2. Stiefer, P.S.; Coe, D.L. Atlas of CRPAQS Ground Truth Surveys; Prepared for the California Air Resources Board and the San Joaquin Valley Air Pollution Study Agency, March 2002.
3. "ArcIMS 3.1: Architecture and Functionality." An ESRI Whitepaper; Environmental Systems Research Institute, July 2001.
4. Vision statement for the Open GIS Consortium, Inc. Available at <<http://www.opengis.org/>>.
5. Dangermond, J. "g.net – A New GIS Architecture for Geographic Information Services", *ArcNews Online*. Available at <<http://www.esri.com/news/arcnews/spring01/articles/gnet.html>>; last accessed March 21, 2002.

KEY WORDS

geographic information systems (GIS), emissions activity, data management

ACKNOWLEDGEMENTS

Funding for this work was provided by the California Air Resources Board. Karen Magliano, Richard Bradley, Beth Schweher, and Robert Effa from the California Air Resources Board have provided invaluable support throughout the application development effort.

DISCLAIMER

The opinions expressed are those of the authors alone.

TABLES

Table 1. Sites selected for the CRPAQS ground truth data collection.

| Site ID | Site Name |
|---------|-----------------------------------|
| BAK | Bakersfield |
| BTI | Bethel Island |
| SWC | Chowchilla |
| COP | Corcoran-Patterson |
| CLD | Crows Landing-Davis |
| DUB | Dublin |
| FED | Feedlot-Dairy |
| FEL | Fellows |
| FRE(3) | Fresno – 3 sites |
| HAN | Hanford – Irwin Street |
| HEL | Helm |
| KCW | Kettleman City |
| LVR | Livermore – Rincon Pine |
| MRM | Merced Midtown |
| M14 | Modesto 14 th Street |
| SDP | Sacramento – Del Paso |
| SJ4 | San Jose – 4 th Street |
| SEL | Selma Airport |
| SNF | Sierra Nevada Foothills |
| SOH | Stockton - Hazelton |
| VCS | Visalia – Church Street |
| WAG | Walnut Grove |

Table 2. Example of emissions activity information contained in the Atlas of the CRPAQS Ground Truth Surveys for the San Jose 4th Street site.

| Site | Sourceid | Sourcetype | Surveydate | Name |
|------|----------|------------------|------------|--------------------------------|
| SJ4 | 1 | Misc. Business | 7/21/2000 | The Printing Press |
| SJ4 | 2 | Construction | 7/21/2000 | Opus West Construction Company |
| SJ4 | 3 | Misc. Business | 7/21/2000 | Flow Master |
| SJ4 | 4 | Misc. Business | 7/21/2000 | Plywood and Lumber Sales |
| SJ4 | 5 | Misc. Business | 7/21/2000 | General Auto |
| SJ4 | 6 | Residential Zone | 7/21/2000 | none |
| SJ4 | 7 | Food Preparation | 7/21/2000 | Marisco's India |
| SJ4 | 8 | Residential Zone | 7/21/2000 | none |
| SJ4 | 9 | Vehicle | 7/21/2000 | Greyhound bus terminal |
| SJ4 | 10 | Food Preparation | 7/21/2000 | Museum Cafe |
| SJ4 | 11 | Food Preparation | 7/21/2000 | Gordon Biersch Brewery |

FIGURES

Figure 1. Depiction of the CRPAQS study domain and monitoring sites selected for the CRPAQS ground-truthing survey.

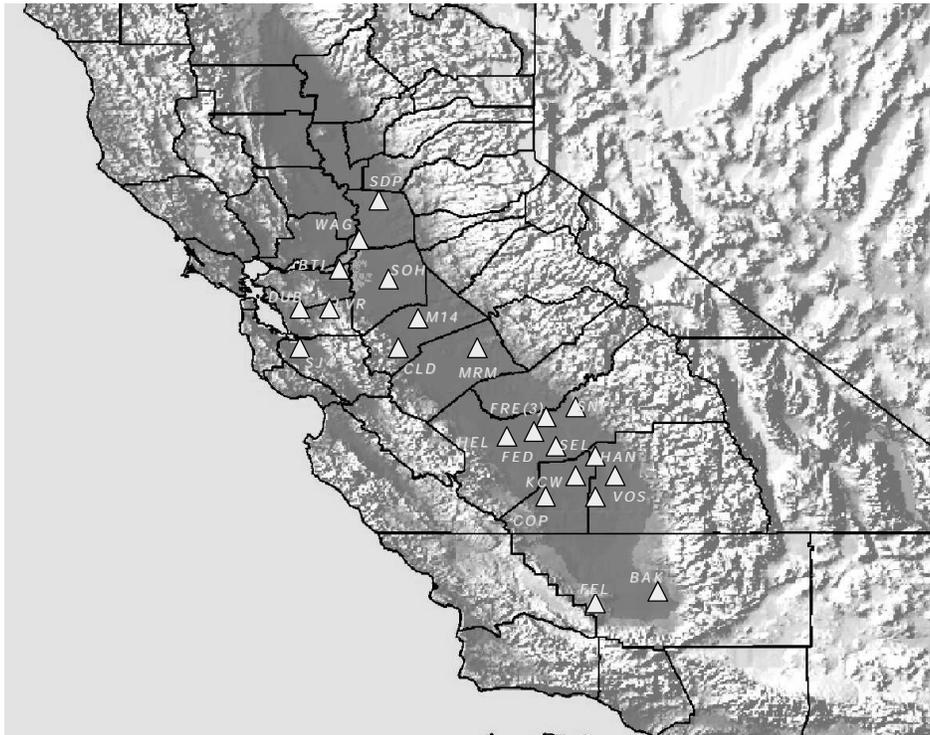


Figure 2. Example of spatial emissions activity information contained in the Atlas of the CRPAQS Ground Truth Surveys for the San Jose 4th Street site.



Figure 3. Schematic of the ArcIMS architecture.

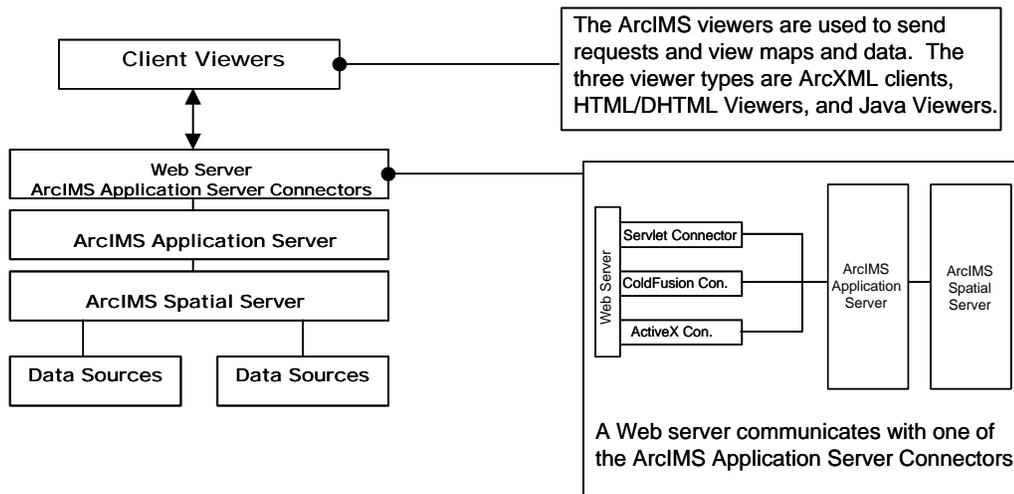


Figure 4. Schematic of the COARSE application and data flow.

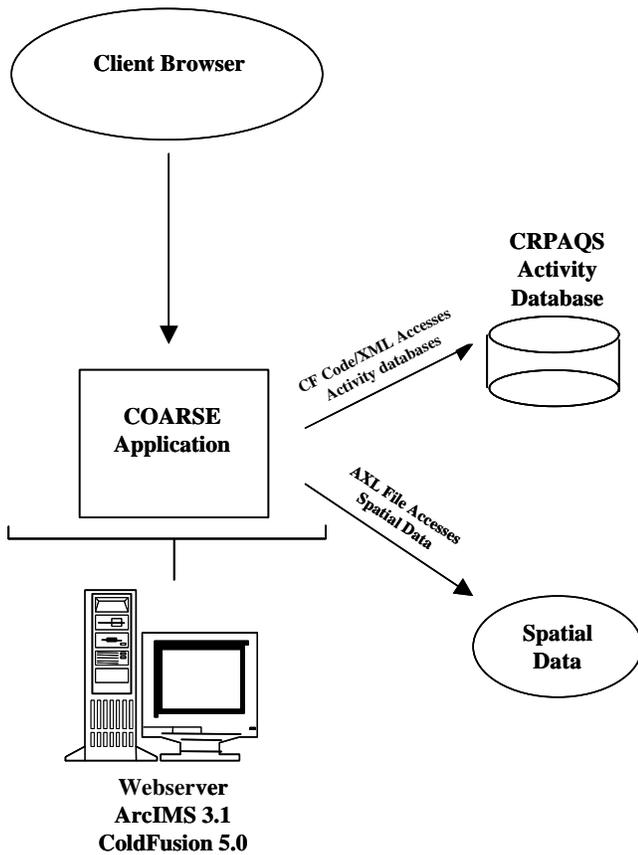


Figure 5. Annotated screen shot of the COARSE user interface, feature, and functions.

The screenshot shows the COARSE Interactive web application in a Microsoft Internet Explorer browser window. The main map displays a grid of monitoring sites (numbered 1-63) in the San Jose area, with labels for streets like N 1st St, N 4th St, and N 10th St, and landmarks like Ryland Park and San Jose State Univ. A scale bar at the bottom of the map indicates 0.5km, 1km, and 2km.

Navigation Tools: Located in the top-left sidebar, it includes a "Zoom to a new:" section with "Site" and "County" tabs, a dropdown menu set to "Angiola", and a "Go" button. Below are icons for "Zoom In", "Zoom Out", "Full Extent", and "Pan".

Activity Reports: Located in the middle-left sidebar, it contains buttons for "Short Report", "Long Report", "Emissions Events", "Angiola Camera", and "Site Logs".

Map Layers: Located in the top-right sidebar, it lists various map features with checkboxes: Monitoring Sites (checked), Shaded Relief (unchecked), Landmarks (checked), Major Roads (checked), Minor Roads (checked), Sources (checked), Railroads (checked), Airports (unchecked), Schools (unchecked), Parks (checked), County Boundaries (checked), Water Bodies (checked), Emissions Events (checked), and Hospitals (unchecked).

Active Layer: Located in the bottom-right sidebar, it features a "Choose layer" dropdown menu and an "Identify" button.

Bottom Navigation: A row of buttons at the bottom includes "Print", "Download Data", "Overview Map", and "Legend".

Annotations: Lines connect text boxes to specific UI elements:

- Navigation Tools:** Points to the top-left sidebar.
- Activity Reports:** Points to the middle-left sidebar.
- Emissions Events (a):** Points to the "Emissions Events" button in the Activity Reports sidebar.
- Angiola Camera (b):** Points to the "Angiola Camera" button in the Activity Reports sidebar.
- Site Logs (c):** Points to the "Site Logs" button in the Activity Reports sidebar.
- Map Layers and Legend:** Points to the Map Layers and Active Layer sections on the right.

Print maps and data files *Download data in text file format* *Overview map toggle* *Legend/map layer toggle*

Site Logs (c)
Retrieves site logs for all sites visible in the map window. The site logs are diaries containing notes and information about emissions events and activities observed by field technicians during the CRPAQS period

Map Layers and Legend
The "Map Layers" list allows users to toggle map layers on and off. Selected map layers contain feature attribute information such as feature names and identifiers. This information can be accessed using the "Active Layer" selection window located in the bottom right section of the interface. When a map layer is selected from the active layer dropdown list, the selected layer becomes active and attribute information about map features can be obtained by clicking on the feature of interest in the map window.