An Interactive Visual Exploration Tool for Northern California's Water Monitoring Network

Jaya Sreevalsan-Nair[a] Erwin Van Nieuwenhuyse[b] Ingrid Hotz[c]
Lars Linsen[d] Bernd Hamann[a]

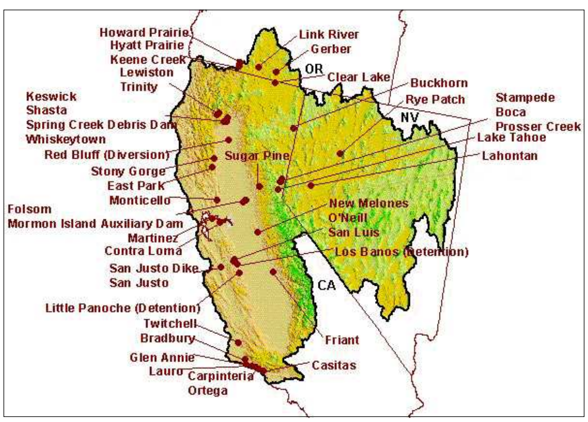
[a]Institute for Data Analysis and Visualization, University of California, Davis, CA 95616, U.S.A.
[b]U.S.Bureau of Reclamation, 2800 Cottage Way, Sacramento, CA 95825, U.S.A.
[c]Visualization & Data Analysis, Konrad-Zuse-Zentrum für Informationstechnik Berlin, Germany
[d]Computational Science and Computer Science, International University of Bremen, Germany

ABSTRACT

The water monitoring network in Northern California provides us with an integrated flow and water-quality dataset of the Sacramento-San Joaquin Delta, the reservoirs, and the two main rivers feeding the Delta, namely the Sacramento and the San Joaquin rivers. Understanding the dynamics and complex interactions among the components of this large water supply system and how they affect the water quality, and ecological conditions for fish and wildlife requires the assimilation of large amounts of data. A multivariate, time series data visualization tool which encompasses various components of the system, in a geographical context, is the most appropriate solution to this challenge. We have developed an abstract representation of the water system, which uses various information visualization techniques, like focus+context techniques, graph representation, 3D glyphs, and colormapping, to visualize time series data of multiple parameters.

Keywords: Information Visualization, Focus+context, Multivariate data, Time series data

1. INTRODUCTION



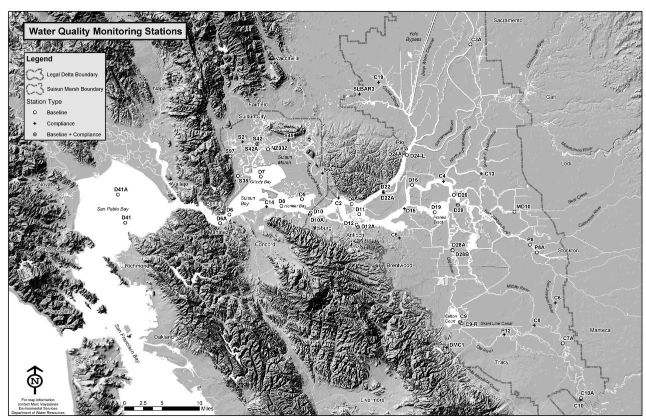


Figure 1. *Left:* Reservoirs/dams in the Mid-Pacific Region. *Right:* Monitoring Stations in the delta of Sacramento and San Joaquin rivers. (Image courtesy of Dr. Randy Dahlgren, U.C.Davis)

The water supply system of Northern California forms a well-knit network of inlet streams, reservoirs, and highly regulated river channels that flow into the delta of Sacramento-San Joaquin rivers and the San Francisco Bay. Figure 1 shows the actual geographical locations of the dams and monitoring sites in the Delta, and Figure 2 shows the schematic layout of the monitoring stations on the Sacramento and San Joaquin rivers. Historically, the Sacramento River floods in early spring and winter, and the San Joaquin River gets flooded in late fall and

Further author information:
E-mail: [a]{jnair,bhamann}@ucdavis.edu, [b]evannieuwenhuyse@mp.usbr.gov, [c]hotz@zib.de, [d]l.linsen@iu-bremen.de

winter. The Central Valley Project of the U. S. Bureau of Reclamation and State Water Project of the California Department of Water Resources were developed to transfer water from the Sacramento River watershed, which receives 67-75% of Northern California's precipitation, to much-drier tracts in the San Joaquin River watershed, which receives 25-33% of regional precipitation. Large pumping plants operated by federal, state and local water supply agencies export water from the Delta via aqueducts to destinations as distant as Southern California. The entire water system supports many beneficial uses including agricultural, municipal and industrial water supply, hydroelectric power generation, waste water and waste heat assimilation, fish and wildlife production and recreation. The developed water is stored in reservoirs and conveyed via river channels, the Delta and the aqueducts. The complex flood-management systems in both the Sacramento and the San Joaquin River basins provide us with a variety of datasets of physical properties and environmental conditions, which is referred to as "water-quality data."

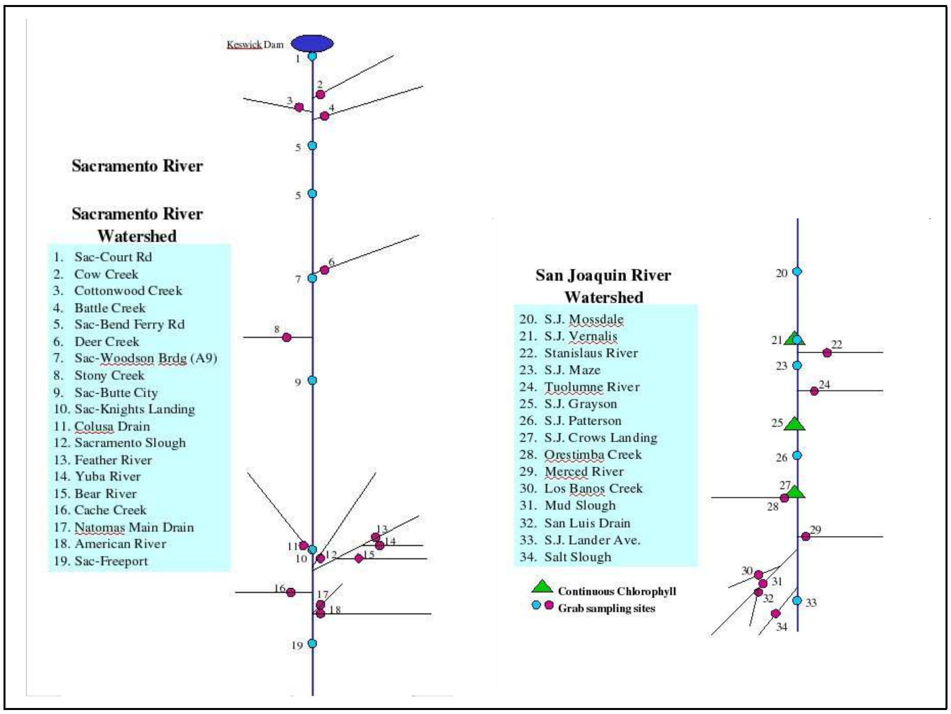


Figure 2. Schematic layout of monitoring stations on *left:* Sacramento River and *right:* San Joaquin River. (Image courtesy of U.S.Bureau of Reclamation)

Several visualization models exist for specific applications in meteorology, hydrology, and other scientific and engineering systems. We have developed a similar paradigm to achieve multivariate, time series data visualization of the rivers, the reservoirs, and other components of the water system in Northern California. Our paradigm is a combination of information visualization and scientific visualization methods, in the sense of iconic representation and focus+context, and representation of multiple parameters.

Information visualization is a process of transforming data that is not inherently spatial into a visual form, allowing the user to analyze and understand information. Information visualization complements scientific visualization that focuses on spatial data generated by scientific processes.[1] Though the visualization community debates over the distinction of these two branches, several applications use a combination of techniques from either methodology. Cartographic information is used in both types of applications.[1] We use a cartographic background in our visualization, for projecting water bodies in the geographical context, similar to the operational weather forecasting model.[2] Other application-driven visualizations include task-based visualization[2] which is used to promote a high-level reuse of interface and content elements for specific user applications and to build a general-purpose toolkit. Pang et al.[3] used a similar concept for a modular visualization environment. Our model, representing the data from the different components, is very similar to the many of these models.

However, our visualization application has some inherent challenges, which include non-uniform timeseries data, and clustered object layout. Maximizing screen-space utilization and placing various components of the system in the right geographical context requires constricted space transformation. Since the different components have different geometry and measured data in different time frames, we have adopted different yet intuitive representations for each of them. For example, the monitoring sites on the rivers has a graph-like structure with links and nodes, unlike the scattered points on the Delta. Due to the specific end-user target for such a tool, we also need specific intuitive representations for various variables, like flow, concentrations of different chemical solvents, etc.

2. PROBLEM DEFINITION

Our goal was to develop a visualization tool for the Northern California water resources, which consists of reservoirs in Mid-Pacific Region, the Sacramento and San Joaquin rivers, and the Sacramento-San Joaquin delta. We treat the flows in the Delta differently from the water-quality data at the monitoring stations, as they are independent and different entities. Further, the components of the system are referred to as reservoirs, rivers, Delta and Delta flow, respectively.

2.1. Data

The latitude-longitude, and flow and water-quality data of the monitoring sites of the rivers and the Delta, and the reservoirs are given. The water-quality data for the rivers and the Delta, which is measured biweekly, consists of the measurements of physiochemical properties of the water at these sites. Water-quality data of the sites on the rivers include temperature, pH, electrical conductivity, turbidity, and the concentration of certain constituents like Na, Ca, K, Mg, Cl, SO_4, HCO_3, Si, NH_4-N, NO_3-N, PO_4-P, chlorophyll, phaeophytin, volatile suspended solids and nonvolatile suspended solids. Water-quality data of the sites on the Delta include temperature, pH, electrical conductivity, turbidity, concentration of total phosporus, phosphates, total nitrogen, dissolved nitrogen, SiO_2, chlorophyll, phaeophytin, volatile suspended solids and total suspended solids. Main daily flow data for the rivers and the Delta are used. Data for the reservoirs consist of daily measurements of volume of the reservoirs, which are given for the period 1999-2006. Flow and water quality of various monitoring sites on the rivers are for the period 1999-2003, and that in the delta are given for the period 1975-2005. In the current application, we use a dataset of 35 river-monitoring sites with flow-data and 27 different parameters in the water-quality data, 44 Delta-monitoring sites with 15 different parameters in the water-quality data, 6 different flows in the Delta, and volume data from 14 reservoirs.

2.2. Requirements of the Visualization Tool

The visualization tool of such a large system, after considering its constraints, should facilitate the following capabilities:

- Display of time series data: The data available for the rivers, the Delta, the Delta flow and the reservoirs are not uniform, i.e., all of them do not have the same date range. But most of them have overlapping periods of measurement. Our visualization model facilitates both joint as well as individual representations of each of the components.
- Display of multiple parameters: Data available for the monitoring stations of the rivers and the Delta is constituted by multiple variables, which our visualization should be able to display as many as possible, simultaneously and without causing clutter. The representation of these variables should be able to capture statistical characteristics of the data, like its time-varying characteristics, extrema, etc. To enable these features, we use a method of representing the variables with a dual context − of global and local contexts. This gives the user a feel of the range of values each variable represents.
- Optimizing screen space and maintaining geographical context: All the components of this system are concentrated largely in a small area in Northern California, in small clusters. To reduce occlusion, it is desirable to zoom into these clusters. However, it is also desirable to maintain the geographical context of the components on the map and their relative positions. To meet these requirements, we use focus+context techniques to achieve uniform spacing between the various components.

- Different representations for various components: The river, the Delta, the Delta flow and the reservoirs are essentially different entities with their own unique geometric definitions. Hence they have to be represented in different yet intuitive ways. The graph-like structure for the rivers calls for an ordered structure unlike the Delta and the reservoirs which, respectively, are scattered point data. The Delta intuitively has a triangular structure, and the reservoirs are analogous to cylinders or tanks with the property of "storage."
- Intuitive representations for different variables: Depending on certain inherent characteristics of different variables, appropriate treatment has to be given to their representations. For example, flow values give a geometric definition of the rivers, variables like temperature and pH has its own intuitive color-schema, namely, blue-to-red and red-to-blue, respectively, and variables like chlorophyll has its own intuitive color tones, namely green. We adopt different strategies like geometric primitives, three-dimensional glyphs, and color mapping to efficiently represent different variables.
- User interaction and flexibility: This is a feature required in any scientific visualization tool which helps the user control the display.

3. GEOMETRIC DEFINITION

3.1. Optimized Layout

The monitoring sites of the rivers and the Delta, and the reservoirs are positioned on the map of California using their latitude and longitude information. It can be observed that they are concentrated in a small region in Northern California, as shown in Figure 3.

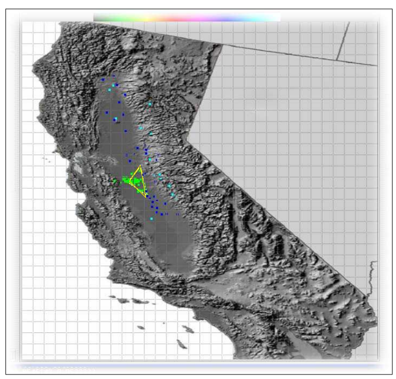


Figure 3. Initial layout of sites. Blue dots are monitoring sites on the rivers, green dots are monitoring sites on the Delta, and cyan dots are the reservoirs; the yellow triangle shows the extent of the Delta.

Focus+context methods are techniques used for "zooming into" relevant data and keeping the rest of the information visible albeit "out of focus." In our case, we use distortion techniques, as our display is static with respect to geographical context. The most prominent "focus+context" techniques include fisheye views,[4] perspective wall,[5] document lens,[6] and nonlinear magnification transformations[7].[8] In our work, we need the effect of continuity between the "focus" and "context" regions, which can be achieved well by using fisheye, document lens, and nonlinear techniques. We use the fisheye technique as linear transformations lead to good results for our specific application. Figure 4 shows few of these techniques which are relevant to our work.

For optimal layout, i.e., to maximize screen space utilization, we have implemented a focus+context technique. In the first pass, we implement fisheye effect to zoom into a rectangular region of interest, as done in fisheye calendar.[4] In the second pass, we implement linear transformations within the region of interest. The two passes are as shown in Figure 5. We effectively ensure continuity between the zoomed-in and zoomed-out regions. Theoretically, the two different passes can be combined to give a single linear-transformation step. However, they are kept separate for the benefit of better understanding of shift of the focus between the two passes.

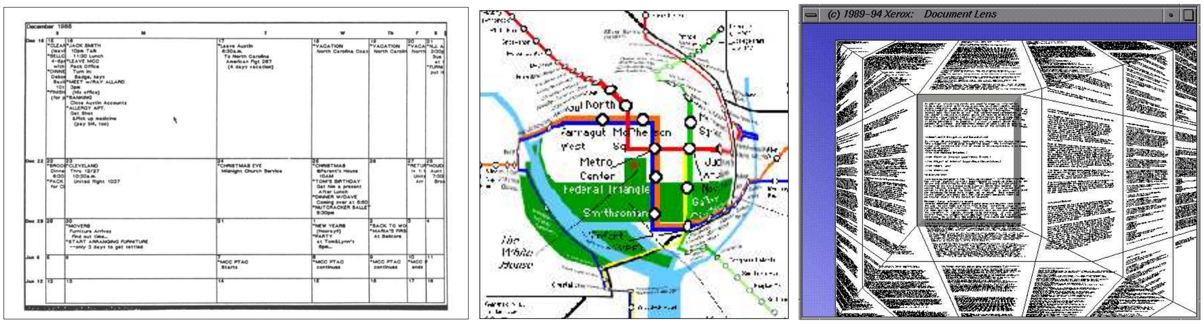


Figure 4. Focus+context techniques. *Left:* Fisheye Calendar, *center:* Nonlinear Magnification, *right:* Document Lens. (Image courtesy of "Generalized Fisheye Views",[4] and of http://www.wikiviz.org/wiki/Overview_plus_detail)

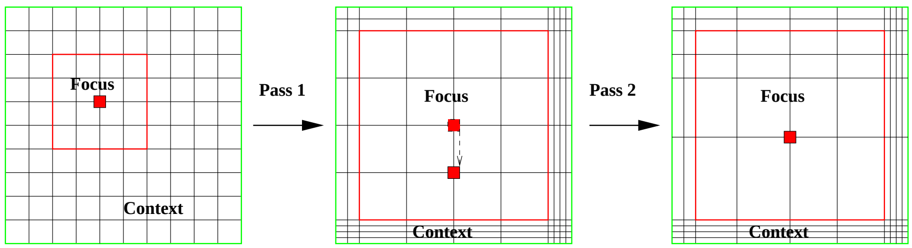


Figure 5. Two-pass transformation for focus+context. The focus shifts between the two passes.

We used a relief map of California in the background for geographical reference. A grayscale map is specifically chosen in order to show the relief features without any background distraction. Our underlying mesh is the rectilinear 2D mesh formed by the latitudes and longitudes in the map. We have represented the mesh in its parametric form, i.e. (s,t) form, in $[0,1]\text{x}[0,1] \in \mathbb{R}^2$ space. Linear transformations are applied on the parameters of the initial mesh to obtain the final mesh. Using the transformation function, new parameters and subsequently, new positions are calculated for any point on the initial mesh. The transformed mesh in (s,t) form also supports parametric texture mapping for the background, as shown in Figure 6. In each dimension, we use four breakpoints to implement piecewise linear interpolation. For the first dimension s, using the breakpoints s_0, s_1, s_2 and s_3, and the new values of the intermediate ones, s'_1 and s'_2, respectively, given that $0 \leq s_0 \leq s_1 \leq s_2 \leq s_3 \leq 1$, and $0 \leq s_0 \leq s'_1 \leq s'_2 \leq s_3 \leq 1$, the transformation function of s to give the transformed value, s', is given as:

$$
\begin{aligned}
s' &= s_0 + \frac{s - s_0}{s_1 - s_0} * (s'_1 - s_0) && \forall s \in [s_0, s_1] \\
s' &= s'_1 + \frac{s - s_1}{s_2 - s_1} * (s'_2 - s'_1) && \forall s \in (s_1, s_2] \\
s' &= s'_2 + \frac{s - s_2}{s_3 - s_2} * (s_3 - s'_2) && \forall s \in (s_2, s_3]
\end{aligned}
$$

Similarly, transformation function in the second dimension, t, is also determined. In our case, we chose appropriate breakpoints in both dimensions, for both the passes, using trial and error method. The chosen values led to the best optimum spacing between the sites, and the reservoirs and we have been able to increase the space usage for object layout on the two-dimensional map of California from approximately 15% (Figure 3) to 80% (Figure 6), after two passes. Once we achieved the distorted version, we retained it, and used a reference (undistorted) map to indicate the current cursor position.

3.2. Representation of River Sites

As shown in Figure 2, the schematic representation of the Sacramento and the San Joaquin rivers is analogous to a graph. The monitoring sites become the nodes, and the connectivity between these nodes forms the edges. To

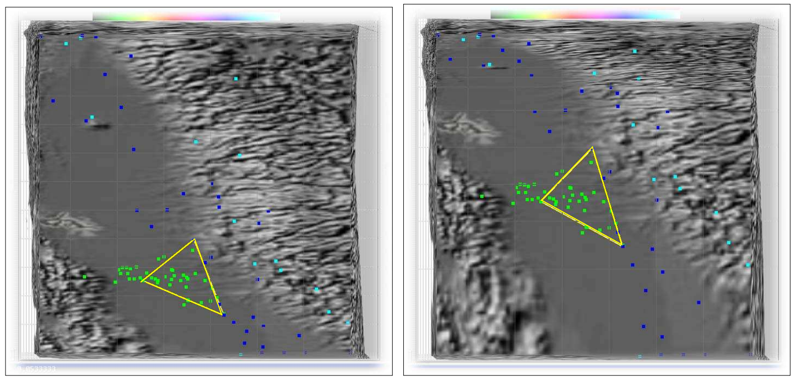


Figure 6. Two-pass space transformation. *Left:* After first pass, *right:* After second pass.

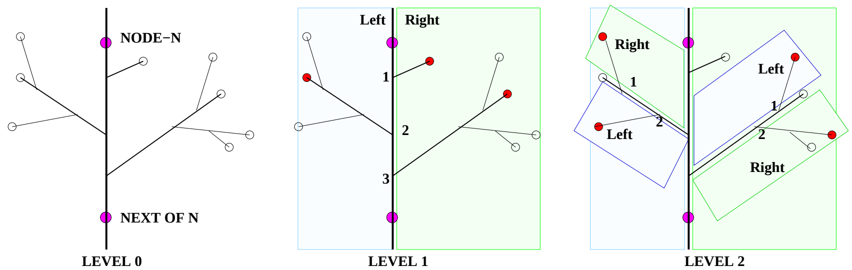


Figure 7. Tree representation of the geometry of the rivers with consistent orientation and consistent ordering. Each node (represented as circles) at each level has a parent, left child, right child, and a next node. Red nodes belong to the indicated level, and the first two levels are shown here.

efficiently represent this structure, we use a tree data-structure which is constructed by implementing a parent-child relationship between consecutive depth levels as shown in Figure 7. As in a tree structure, each node has a left and a right child. In our case, level 0 consists of nodes in the Sacramento and San Joaquin rivers, and are ordered in a linked-list form, where each node has a "next" node, which is the node downstream of it . The children of a node are the upstream-most nodes on the tributaries between the node and its next node. In higher levels, the children of a tributary are the sub-tributaries that enter the tributary. This structure tallies well with the schematic diagram of Sacramento-San Joaquin rivers(Figure 2). Also, we follow a consistent left-right orientation and numerical ordering of sites in a branch, again in the downstream direction, as can be seen in the figure 7.

3.3. Representation of Delta

The Delta has a different representation compared to the rivers, despite similar data types, due to strong local variations in the data of the Delta. There are no continuous functions to define any of the variables of the Delta. The three inflows to and the three outflows from the Delta, that constitute the flows in the Delta, do not have a connectivity structure as that of the rivers. Hence the different flows are depicted as discrete rectangles (to depict the flow width) going in or out of the Delta, which is intuitively represented as a triangle. The other variables of the Delta are represented discretely using traditional techniques of visualization of scattered point datasets. The representation of the flows and the variables are further discussed in Section 4.

3.4. Representation of Reservoirs

The reservoirs are storage components unlike the sites on the rivers and the Delta, which are in transit. Hence they are represented as cylinders which are analogous to tanks. However, just as the sites in the Delta, the reservoirs are discrete entities and are also represented as scattered point datasets. The radius and height of the cylinders are determined by the ratio of the capacity of each of the reservoir to the maximum capacity possible, to depict the relative capacity of the reservoirs.

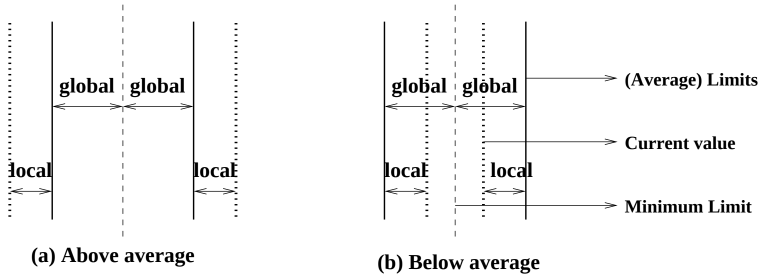


Figure 8. Global-local scaling. Representations of values above and below average

4. MULTIVARIATE DATA VISUALIZATION

For all the different representations of the various variables known for the rivers, the Delta, the Delta flow or the reservoirs used in this system, we have used a dual approach for representing global as well as local statistics. A basic structure is derived by scaling using the global values and is used to indicate relative behavior in a global context. Dynamic surfaces or outlines are then used to indicate the local variations.

4.1. Interpolation and Scaling of Data

Owing to the connective nature of the monitoring sites on the rivers, the variables like flow and fluxes of parameters can be linearly interpolated. However in the case of the Delta, the Delta flow and the reservoirs, the sites are discrete. Hence, data cannot be interpolated in the latter cases.

For implementing our dual approach of global-local variations, the global and local scaling ratios, r_{global} and r_{local}, respectively, are determined for each variable of each entity, namely the sites, the flows, and the reservoirs. To find r_{global} of any site or flow, the average values of each of the entities are scaled with respect to the maximum of the average values of all the entities in that category (say, site or flow), respectively. Maximum of averages are used instead of the absolute maximum because the average value gives a more uniform distribution compared to the maximum value (which occurs very rarely). Hence for local scaling, we use the averages, and in global scaling, we use the maximum of the averages. Except in the case of reservoirs, r_{global} is determined by scaling the capacity of the reservoir with respect to the maximum of all capacities. To find r_{local} of a site, a flow or a reservoir, we scaled the current value with respect to the average value of the variable at the site. The corresponding r_{global}'s are used to determine the flow widths for the rivers and the Delta flow, and the dimensions of the cylinder for the reservoirs. The corresponding r_{local}'s are used to indicate the measure by which the dynamic outlines or surfaces displaced, which in turn indicates the temporal changes in the data. Figure 8 demonstrates the idea. In cases of scaling current values with respect to the respective average value, r_{local} was calculated as a piecewise linear function using the corresponding minimum, average and maximum values, given by:

$$r_{local} = \begin{cases} -0.5 + 0.5 * \frac{\text{current\_value-minimum}}{\text{average-minimum}} & \forall \text{ current\_value} \leq \text{average} \\ 0.5 * \frac{\text{current\_value-average}}{\text{maximum-average}} & \forall \text{ current\_value} > \text{average} \end{cases}$$

This gives the following ranges: $r_{global} \in (0, 1]$ and $r_{local} \in [0, 1]$, if r_{global} is scaled with respect to maximum, otherwise, $r_{local} \in [-0.5, 0.5]$, which also ensures that the average value corresponds to $r_{local} = 0.0$.

4.2. Flow Representation for Delta

Three inflows to and three outflows from the Delta constitute the major flows in the Delta. The inflows are SAC, YOLO and SJR, which are inflows from the Sacramento River from the Sac-Freeport Station, from the Yolo Bypass, and from the San Joaquin River, respectively. The outflows are CVP, SWP, and OUT, which are outflows to the Central Valley Project, to the State Water Project, and the net outflow from the Delta to the San Francisco Bay.

The current flows are represented as rectangles, placed along the perimeter of the triangle representing the Delta. The positions of the flows are contextual to their actual geographical positions. The widths of the

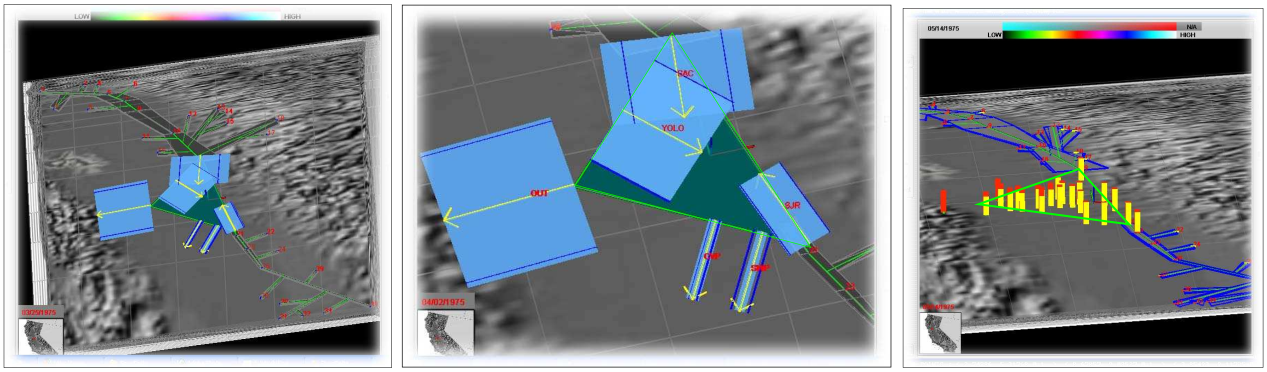


Figure 9. *Left:* Delta flow representation; *center:* close-up of Delta flow; *right :* Columns to represent parameters at sites on the Delta: Yellow columns show temperature, and red columns show pH.

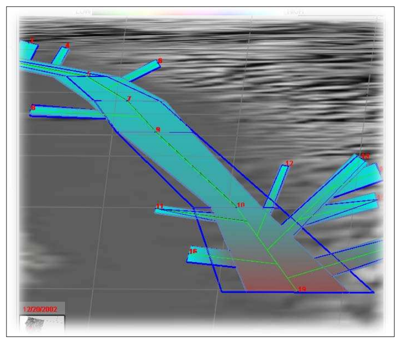


Figure 10. Representation of river flow as trapezoids with time-varying outlines and color mapping of an additional parameter: Temperature color-mapped on Sacramento River using blue-to-red scheme.

rectangles have two referential limits, on either side of its axis of symmetry (which is normal to the triangle sides), which indicate the average of the flow values, and are referred to as "average-limits" of the respective component flow. The distance between the limits of a component flow is determined using the r_{global} of its average flow value. The average-limits of the component flow are constant, and the current flow is represented by a rectangle whose width is determined using the r_{local} of the flow values. Thus, the time series representation of the flow data is done by the flow rectangles whose areas move "inwards" or "outwards" the average-limits, to depict below-average or above-average flows, respectively. Figure 9 shows a snapshot of the Delta flow.

4.3. Parameter Representation for Delta

The water-quality data of the monitoring sites on the Delta are represented discretely as columns, as shown in Figure 9. The length of the column is directly proportional to the measured value, and they are scaled linearly and globally.

4.4. Flow Representation for Rivers

The rivers are given a basic "flow-structure" using a chain of trapezoids, which gives an intuitive geometric representation of flow. The widths of the trapezoids are determined using r_{global} value of flow value at each site. The monitoring sites are the midpoints of the parallel pair of sides of the trapezoid. The chain is constructed using the tree structure of the rivers (see Figure 7). The current flow is represented as outlines of trapezoids moving in and out of the basic structure, indicating below-average and above-average flows, respectively, as shown in Figure 8. The widths of the dynamic trapezoids are determined using r_{local} values of the flow. Figure 10 shows the flow structure of the Sacramento and the San Joaquin rivers. An additional variable can be represented by color filling the trapezoids. The color is determined by using a user-selected transfer function. The color scheme

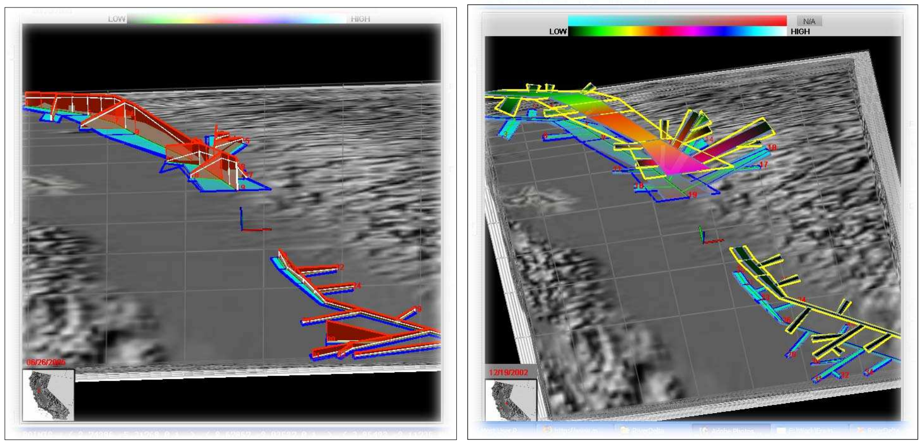


Figure 11. Parameter flux representation in rivers (Sacramento River in the North and San Joaquin River in the South). Temperature flux represented using *left:* Using vertical surfaces; *right:* using horizontal surfaces.

for the transfer function can be selected intuitively for each variable, for example, blue-to-red for temperature, red-to-blue for pH, green-toned for chlorophyll, etc.

4.5. Parameter Representation for Rivers

Our tool facilitates three different representations of parameters at the sites on the rivers. Their fluxes are calculated as product of concentration of the parameter and the flow at the site. The different representations, as shown in Figures 11 and 12, include:

- Vertical surfaces for representing fluxes: In this representation, the vertical surface is a chain of trapezoids constructed from vertical lines at each site. The lengths of these vertical lines are determined using r_{local} with respect to the average flux values. A white line represents a constant reference structure which is constructed by joining the ends of another set of vertical lines from each site. The lengths of these vertical lines are determined using r_{global} of the average flux value of the parameter at the site. Just as in the flow representation, the white line is similar to the basic "flow-structure." The color-filled trapezoids move below and above the white-line to indicate below-average and above-average values, respectively. The fill-color of the trapezoids is user-selected and is a constant map for a parameter.
- Horizontal surfaces for representing fluxes: In this representation, the basic flow-structure is displaced in positive z-axis and the dynamic outlines, in this case, show the variation in the flux of the chosen parameter. The outlines are computed using r_{local} with respect to the average flux values. The trapezoids are color-filled and colormapping in this case is determined using a transfer function based on the r_{global} values. The transfer function is determined from a color-map.
- Columns for representing concentration: As in the case of the delta sites, the water quality data of the river sites are represented using columns whose heights are directly proportional to the measured values.

4.6. Volume Representation for Reservoirs

As explained in Section 3, cylinders are used to represent reservoirs. The reservoirs have only one parameter to be displayed, namely, the current volume of water. The fluctuating height indicated by the opaqueness in the cylinder represents the current "filling-in" or "draining-out" behavior of the reservoir.

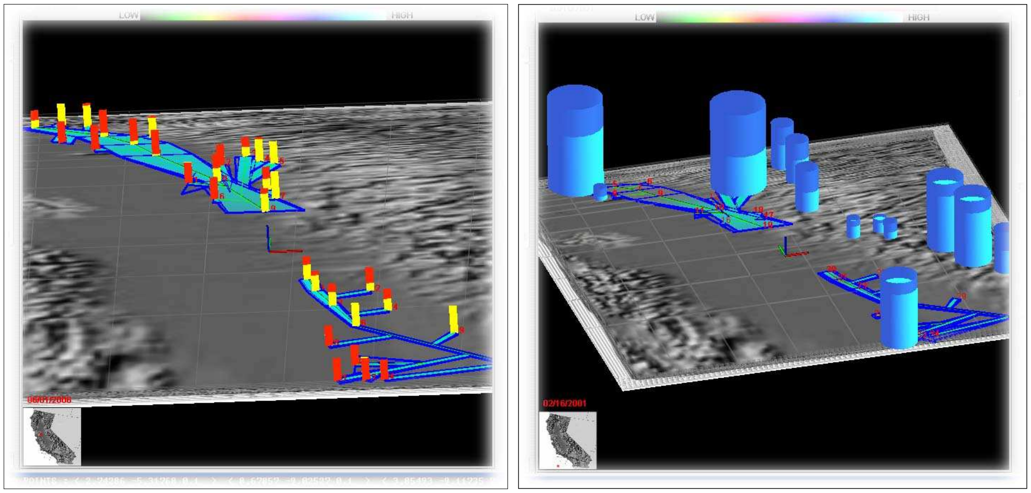


Figure 12. *Left:* Columns to represent parameters at river sites on Sacramento (in the North) and San Joaquin (in the South) Rivers; *right:* cylindrical representation of water volume in reservoirs

4.7. Previous work

Summarizing, our approach uses a mixed model of both information and scientific visualization techniques. Similar work has been done in different applications with similar requirements. Using three-dimensional glyphs to represent the water bodies is similar to the representation of nodes in time-oriented geographical networks.[9] Kolojejchick et al.[10] used techniques to create a suite of basic tools and specialized information "appliances" by combining basic analysis and reporting tools into an integrated information workspace. Several scientific visualization models,[11–13] web-based applications,[14] and visualization architectures,[15] have been developed for similar applications, in which either scientific visualization or traditional information visualization techniques, like scatter plots, color maps, etc., are used. An earlier version of our work[16] focussed on obtaining a multiresolution model.

5. INTERFACE

Our visualization tool (see Figure 13) is both user-interactive as well as flexible to a certain extent to fulfill the end-users' needs. The following features allow user interaction :

- Selection of components to be displayed: The user can decide which components, namely, the rivers, the Delta, the Delta flow, and the reservoirs, need to be displayed. This facilitates independent as well as joint display of the various components.
- Selection of variables to be displayed: The user can select and simultaneously display different variables for the various components. The tool allows a maximum of two variables to be displayed simultaneously for the rivers as well as for the Delta, in addition to an extra variable that is colormapped on the "flow structure" of the rivers. The limit on the number of variables is motivated by the application as multivariate visualization enables study of interdependence of various variables, and the application of our tool usually studies interdependence of at the most three variables. However, our tool is flexible to be extended to accomodate simultaneous visualization of more variables until occlusion occurs. In the column-display, one can extend the tool upto 5 variables, and for horizontal or vertical displays of variables, one can view upto 3 variables at a time.
- Selection of color schema: The user can select the color schema to represent the variables. This helps them to intuitively visualize the parameters in its characteristic colors. The user can choose the color for the each of the parameters of the river as well as the delta. The user can also select color-maps for the colormapping of an additional variable within the "flow structure" of the river.

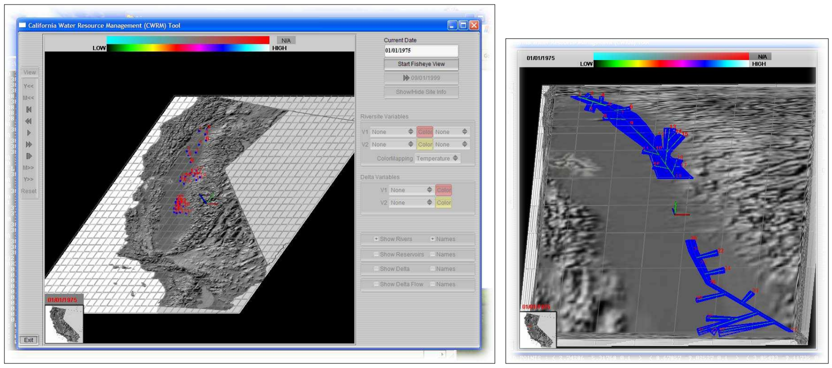


Figure 13. *Left:* A snapshot of the complete visualization tool; *right:* the transformed cartographic background and the flow structure of Sacramento (North) and San Joaquin (South) rivers after applying distortion focus+context techniques.

- Data addition and display: The data files can be easily accessed and new readings can be added. The duration of the time frame being displayed by the tool modifies depending on the earliest and latest dates of the measurements of all the components. Textual display of data is an additional feature to enable the user to see the absolute figures of the data on mouse-over a site or a reservoir. The user can pan and rotate using the right- and left-mouse buttons, respectively, and zooming is enable using keyboard controls.

- Playback features: These enable control in the time series visualization. They enable the user to march through time without any interaction in a loop, march ahead or behind time by a day, week, month or year.

6. CONCLUSIONS

We have developed an integrated user-interactive visualization system (see Figure 13), which allows scientists to study a significant part of the complete hydrological system in California. The multivariate time series data visualization helps the analysis of such a large dataset by simultaneous display of time series at multiple locations and by facilitating various combinations of parameters. We also have successfully optimized the layout of the various components in the system to utilize maximum screen space. We have different representations for flow and water quality data for the rivers and the Delta. For the rivers, we have two different representations for the fluxes of the parameters, of which the vertical surfaces are good for visualizing multiple parameters as they allow occlusion-free comparisons. The horizontal surfaces are appropriate for comparing variations in a parameter with respect to the variations in flow. The use of focus+context techniques supports an optimum layout of objects for maximizing the utilization of screen space. This is an apt technique to be used with cartographic references, as the planar latitude-longitude maps are akin to 2D rectilinear meshes. Summarizing, the applications of our visualization tool are as follows:

- Joint visualization: The visualization enables the user to visualize all components of the water system, i.e., the rivers, the Delta and the reservoirs. This allows the user to get a good insight of hydrological cycles, seasonal cycles, water "transactions" between the Sacramento and San Joaquin watersheds, etc.

- Multivariate time-series data visualization: Since we are visualizing a water-monitoring system, the visualization system helps us to obtain understanding of chronology of the beginning of measurement at various sites in the system, drastic climactic periods, etc. The tool also allows the user to study interdependence of various parameters among all the components of the system.

- Occlusion-free layout of the components: Use of focus+context methods enables us to derive an optimal layout of all the sites and the reservoirs. This helps us to have a occlusion-free structure to depict the rivers and fairly uniform spacing between the sites and the reservoirs. We use predefined geometric primitives to build the structures for the rivers, the delta, and the reservoirs.

- Interactive tool: The tool allows the user to select the components, variables, and the color schema touse. It also has playback features to navigate through the time series data, apart from navigation control features like rotation, panning, and zooming.

The visualization tool we have presented is simple yet powerful for analyzing data for scientists studying the hydrology, ecology, water quality, marine sciences, and other scientific studies of the watershed regions in Northern California. We can extend our tool to include more components and more parameters. Working with large datasets for such a multi-faceted system like this, a multiresolution or hierarchical approach is a solution.

ACKNOWLEDGMENTS

The authors would like to thank the U.S. Department of the Interior, Bureau of Reclamation, Mid-Pacific Region, for the financial support provided for this project (Project# 02FC200126). For the river sites, the water-quality data is courtesy of Dr. Randy Dahlgren, Department of Land, Air, and Water Resources, U.C.Davis, and the flow data is courtesy of U.S. Geological Survey (http://waterdata.usgs.gov/nwis). For the Delta sites, the water-quality data is courtesy of the Environmental Monitoring Program (http://www.iep.ca.gov/emp), and the flow data is from the DAYFLOW model (http://www.iep.ca.gov/dayflow/output).

REFERENCES

1. T.-M. Rhyne, "Does the Difference between Information and Scientific Visualization Really Matter?," *IEEE Comput. Graph. Appl.* **23**(3), pp. 6–8, 2003.
2. L. A. Treinish, "Task-specific Visualization Design: A Case Study in Operational Weather Forecasting," in *Proceedings of the conference on Visualization '98*, H. H. David Ebert and H. Rushmeier, eds., pp. 405–409, IEEE Computer Society Press, 1998.
3. A. Pang, C. M. Wittenbrink, and T. Goodman, "CSpray: A Collaborative Scientific Visualization Application," in *Proceedings of the 1995 SPIE Conference on Multimedia Computing and Networking (MMCN 1995)*, J. M. A. A. Rodriguez, ed., pp. 317–326, ACM SIGMM/SPIE, SPIE, (San Jose, California), March 1995.
4. G. W. Furnas, "Generalized fisheye views," in *Proceedings of CHI-86*, M. Mantei and P.Orbeton, eds., pp. 16–23, (Boston, MA), 1986.
5. J. D. Mackinlay, G.G.Robertson, and S.K.Card, "The perspective wall: detail and context smoothly integrated," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, G. M. O. S.P.Robertson and J. S.Olson, eds., pp. 173–176, ACM Press, 1991.
6. G. G. Robertson and J. D. Mackinlay, "The document lens," in *Proc. of the 6th Annual Symposium on User Interface Software and Technology (UIST'93)*, R. Pausch, ed., pp. 101–108, (Atlanta, GA), 1993.
7. Y. K. Leung and M. D. Apperley, "A review and taxonomy of distortion-oriented presentation techniques," *ACM Trans. Comput.-Hum. Interact.* **1**(2), pp. 126–160, 1994.
8. T. A. Keahey and E. L. Robertson, "Nonlinear magnification fields," in *INFOVIS '97: Proceedings of the 1997 IEEE Symposium on Information Visualization (InfoVis '97)*, J. Dill and N. Gershon, eds., p. 51, IEEE Computer Society, (Washington, DC, USA), 1997.
9. K. C. Cox, S. G. Eick, and T. He, "3D geographic network displays," *SIGMOD Record* **25**(4), pp. 50–54, 1996.
10. J. Kolojejchick, S. F. Roth, and P. Lucas, "Information Appliances and Tools in Visage," *IEEE Comput. Graph. Appl.* **17**(4), pp. 32–41, 1997.
11. R. Stein, A. M. Shih, M. P. Baker, C. F. Cerco, and M. R. Noel, "Scientific visualization of water quality in the Chesapeake Bay," in *IEEE Visualization*, B. H. Thomas Ertl and A. Varshney, eds., pp. 509–512, 2000.
12. S. Thorpe, J. Ambrosiano, C. Coats, A. Eyth, S. Fine, D. Hils, T. Smith, A. Trayanov, R. Balay, and M. Vouk, "The Package for Analysis and Visualization of Environmental Data," in *Computing in Environmental Resource Management*, Air and Waste Management Association, 1996.
13. A. Mankofsky, E. P. Szuszczewicz, P. Blanchard, C. Goodrich, D. McNabb, R. Kulkarni, and D. Kamins, "Case Study: Visualization and Data Analysis in Space and Atmospheric Science," in *Proceedings of the conference on Visualization '94*, R. D. Bergeron and A. E. Kaufman, eds., pp. 341–344, IEEE Computer Society Press, 1994.
14. J. Wood, K. Brodlie, and H. Wright, "Visualization over the World Wide Web and its Application to Environmental Data," in *Proceedings of the 7th conference on Visualization '96*, R. Yagel and G. M. Nielson, eds., pp. 81–ff., IEEE Computer Society Press, 1996.
15. B. Lucas, G. D. Abram, N. S. Collins, D. A. Epstein, D. L. Gresh, and K. P. McAuliffe, "An Architecture for a Scientific Visualization System," in *Proceedings of the 3rd conference on Visualization '92*, A. E. Kaufman and G. M. Nielson, eds., pp. 107–114, IEEE Computer Society Press, 1992.
16. J. Sreevalsan-Nair, C. S. Co, E. V. Nieuwenhuyse, L. Linsen, and B. Hamann, "Visualization of water resource data," in *Proceedings of the 2003 UC Davis Student Workshop on Computing, TR CSE-2003-24*, D. Copsey, ed., pp. 28–29, 2003.