Journal of Floodplain Management
Floodplain Management Association
JULY 2000 Vol. 2, No. 1

Development of Data Base and Statistical Analysis of Water Quality Data for the Los Angeles Region, California

T.V. Hromadka II [1]
D.J. Smith [2]
C.C. Yen [3]

Abstract

The Regional Water Quality Control Board (RWQCB), Los Angeles Region completed a comprehensive update to its 1975 Water Quality Control Plans (Basin Plan) in 1994. These plans, required under both state and federal law, are the cornerstone of water quality regulations for state water quality control agencies and contain water quality standards (comprised of beneficial uses and water quality objectives) for all waters (surface and ground). The purpose of this study was to provide statistical information for the RWQCB, Los Angeles Region, to review, and revise if necessary, water quality objectives for regional rivers and streams as part of a major Basin Plan update in 1994. Additionally, excursion analyses were calculated to assess exceedences of existing water quality objectives. This information was used to perform water quality assessments, which are also required under the Clean Water Act. As part of this study, some four million water quality grab sample measurements, representing 1500 measurement sites, were complied and synthesized into manageable databases. In this paper, this database is reviewed and the software development discussed. These data are available from the RWQCB for further research use in analyzing water quality effects with respect to hydrological processes.

Introduction

Surface water quality monitoring data, collected by the California Regional Water Quality Control Board. California Department of Water Resources, U.S. Geological Survey, and various city and county water resources departments was compiled into a database by the Regional Water Quality Control Board (RWQCB), Los Angeles Region. This database represents data collected at 1500 sites with varying lengths of record, and was tabulated according to various formats.

In order to analyze the data systematically, a computer data base management model was prepared to perform the following tasks:

---

[1] *Professor, Department of Mathematics, California State University, Fullerton, CA*

[2] *Assistant Executive Officer, California Regional Water Quality Control Board, Los Angeles Region, Los Angeles, CA*

[3] *Associate Engineer, Exponent/Failure Analysis Associates, Costa Mesa, CA*

1. Convert each of the raw data files into a uniform and readily manageable data file format.
2. Substitute "non-detect" values into the data file in place of a "non-detect" designation.
3. Separate data according to station, and provide basic statistical information for each constituent for each station.
4. Perform an excursion analysis (i.e., estimate proportion of the data that exceed water quality objectives), per water quality objectives contained in the RWQCB, Los Angeles Region, Water Quality Control Plans (Basin Plans).
5. Prepare trend analysis plots of each constituent.

In this paper, the steps described above will be discussed, and sufficient background information will be presented in order to describe the usefulness of this water quality database for future hydrologic study.

Data Base Development

Data Collection

Water quality monitoring data were compiled by the RWQCB, Los Angeles Region, from various agencies within the jurisdiction of the Los Angeles Region. In general, each data file contains the following information: station identification number, date of sample collection, water temperature, pH, dissolved oxygen (DO), total dissolved solids (TDS), sulfate (SO4), chloride, boron, nitrate, nitrate-N, nitrite N, and phosphate (PO4). The ASCII database files were converted to be compatible with a computer data base management model to be used in the statistical analysis and excursion analysis.

"Non-Detect" Values

Non-detect values in the surface-water database were treated as follows:

1. If the laboratory detection limit was known, one half of the value of the detection limit was used.
2. If the detection limit was not known, other data for the same parameter and time period were examined. If they were similar, this information was used to estimate a limit.
3. If a detection limit could not be estimated, a zero was substituted.
4. If the individual Agency had already substituted a zero for non-detects, the zero was retained.

Some of these detection limits, which were provided by the RWQCB, Los Angeles Region, are listed in Table 1.

Table 1. Known Detection Limits

Constituent	Source Agency	Time Period	Detection Limit
Nitrate ($N0_3$-N)	RWQCB		0.2 ppm
	Department	1978	1.0 ppm
	of Power and	1979	0.9 ppm
	Water (DPW)	1988	0.1 ppm
			0.11 ppm
		1989	0.23 ppm
		1990	0.23 ppm
	City of Pasadena		0.5 ppm
Nitrate ($N0_3$)	RWQCB		1.0 ppm
	DPW	1977	0.1 ppm
		1978	1.0 ppm
		1979	0.9 ppm
	Department of Water Resources (DWR)	(recent)	0.1 ppm
Nitrite ($N0_2$-N)	RWQCB		0.01 ppm
	DPW	1981	0.01 ppm
		1982	0.01 ppm
		1983	0.01 ppm
		1984	0.02 ppm
		1986	0.05 ppm
		1988	1.0 ppm
			0.3 ppm
			0.5 ppm
		1989	0.1 ppm
		1990	0.1 ppm
		1991	0.03 ppm
		1992	0.03 ppm
Phosphate ($P0_4$)	RWQCB		0.01 ppm
	DPW	1990	0.02 ppm
		1992	0.05 ppm
Boron	RWQCB		0.1 ppm
	DPW		0.25 ppm
	Santa Paula Water		0.1 ppm
	Ventura Flood Control puts in a "zero"		

Sorting Data on a Station Basis

After "non-detect" entries in the data base file were replaced by the detection limits (as described in the previous section), the entire data base file contained only numerical values. The next step accomplished by the data management model was to sort the database file on a station basis, for each watershed.

Sorting Station Data on a Constituent Basis

Finally, the station data were sorted into individual data files on a constituent basis. Each resulting data file was named according to the following format: "ABCDEFG.TXT". The first two characters, AB, represent the watershed identification. The next three characters, CDE, represent the station identification, and the last two characters, FG, represent the constituent identification. For example, LA00305.TXT contains TDS values of Station 003 for the Los Angeles River and Tributaries. Table 2 lists the file name convention for each individual file created by the computer data base management model. It is noted that nitrate plus nitrate-N data are used, as there is a combined objective for these two constituents in the Basin Plan.

Table 2. File Name Conventions

Characters		Descriptions
AB		Watershed
	MV	Misc. Coastal Streams- Ventura County
	VE	Ventura River and Tributaries
	SC	Santa Clara River and Tributaries
	CA	Calleguas Creek and Tributaries
	MA	Malibu Creek and Tributaries
	BA	Ballona Creek and Tributaries
	ML	Misc. Coastal Streams/Storm Drains – Los Angeles County
	DO	Dominguez Channel
	LA	Los Angeles River and Tributaries
	SG	San Gabriel River and Tributaries
	SA	San Antonio Creek
CDE		3 digit station identification
FG		2 digit constituent identification
	01	Flow rate (reserved)
	02	Water Temperature (reserved)
	03	pH – Lab
	04	DO
	05	TDS
	06	Sulfate
	07	Chloride
	08	Boron
	09	Nitrate
	10	Nitrate-N
	11	Nitrite-N
	12	$P0_4$
	13	pH – Field
	14	Nitrate-N plus Nitrite-N

Resulting Data Base from Data-Base Management Model

The resulting database developed by the computer database management model includes all the individual data files for each constituent. Each individual data file consists of two columns of data. The first column contains a 6 digit numerical number, which represents the sample collection date. The second column contains the sample value. Table 3 illustrates the data structure for an individual data file. The resulting database was used to calculate basic statistical information and to perform excursion analysis, trend analysis plots, and box-and-whisker plots.

Table 3. Structure of an Example Data Base File

520408	7.20
520408	7.10
530114	8.00
530227	8.10
530407	7.60
560127	7.00
610214	7.80
610214	7.50
620209	7.70
631120	7.60
670131	7.90
701129	7.50
750202	6.70
750919	7.70
780602	7.80
790412	7.70
800616	7.70

Statistical Analysis Procedures

Mean

The sample mean value is defined as the average value of all the samples values. The formula used to calculate the sample mean is:

$$\text{Sample Mean Value} = \frac{\text{(sum of all the sample values)}}{\text{(total number of samples)}}$$

Median

If a sample set is rearranged in the sequence of increasing values, then the rearranged sample set is called the order statistics. The sample median represents the middle value of these order statistics. For an odd number sample size, the middle order statistic is the sample median. For an even number sample size, the median is typically defined to be the average value of the two middle order statistics.

Standard Deviation

The sample standard deviation indicates how the sample values scatter about the sample mean. The sample standard deviation is computed as:

$$\text{Sample Standard Deviation} = \left(\sum_{i=1}^{n} \frac{(X_i - \text{sample mean})^2}{n - 1} \right)$$

Where X_i is the sample value, and n is the sample size.

Percentiles

For a large sample size, the pth percentile is a value such that about p percent of the sample values fall below this value. For example, the value defined by the 70^{th} percentile exceeds about 70% of the sample values (and about 30% of the sample values exceed the 70^{th} percentile). The formula to calculate the 100^{th} percentile is given by Hogg and Tanis (1977) as:

$$\text{pth percentile of the sample} = \begin{cases} \text{The rth order statistic; if r is an integer} \\ \text{The weighted average of the two adjacent order statistics (closest to r); if r is not an integer} \end{cases}$$

where r is defined as (n+1)p.

In this study, when the sample size is too small to estimate the 100^{th} percentile using the previous formula, the maximum value of the sample set is used.

Maximum and Minimum Values of Data Set

It should be noted that the maximum and minimum values of a data set only represent the relative maximum and minimum values of that sample. For example, a sample set consisting of 20 sample points from the uniform distribution, U[0.1], could be (.3407, .144, .696, .8675, .5649, .5793, .1514, .5044, .9859, .4658, .7779, .7986, .052, .6697, .0045, .4999, .493, .7408, .7551, .3124). The maximum and minimum values from that data set are .9859 and .0045, respectively. But the absolute maximum and minimum values are 1. and 0., respectively. From the above, it is seen that the maximum and minimum values are sampling data, and do not match total population values.

Table 4. Established Water Quality Objectives (at the time of this study)

Watershed/name of reach	TDS	Sulfate	Chloride	Boron	Nitrogen	SAR	Stations
Miscellaneous Ventura Tributaries	none						all MV
Ventura River Watershed:							
Above Camino Dielo Road	700	300	50	1.0	5	5.0	VE 30-90
Between Camino Cielo Road and Casitas Vista Road	800	300	50	1.5	5	5.0	VE 40-69
Between Casitas Vista Road and Oak View STP	1,000	300	60	1.5	5	5.0	VE 37-38
Between Oak View STP to Main Street	1,500	600	60	1.5	10	5.0	VE 10-36
Below Main Street	none						VE 5
Santa Clara River:							
Above Lang	600	100	50	.5	5	5.0	SC 850-860
Between Lang and West Pier Hwy. 99	1200	450	100	1.5	10	5.0	SC 760-755
Between W. Pier Hwy .99 and LA/Ventura County Line	1200	550	100	1.5	5	5.0	SC 715-755
Between LA/Ventura County Line And A Street, Fillmore	1300	600	100	1.5	5	5.0	SC 360-710 (except 535-665)
Between A Street, Fillmore and UWCD Diversion Dam near Saticoy (Freeman Dam)	1300	650	80	1.5	5	5.0	SC 140-355 (except 315-350 and 205-220)
Below UWCD Diversion Dam Near Saticoy	none						SC 0-135
Santa Paula Creek above SP Water Works Diversion Dam	600	250	45	1.0	5	5.0	SC 205-220
Sespe Creek above gauging station, 500' downstream from Little Sespe Creek	800	320	60	1.5	5	5.0	SC 315-350
Piru Creek above gauging station Below Felicia Dam	800	400	60	1.0	5	5.0	SC 535-665
Calleguas Creek:							
Above Potrero Road	850	250	150	2.0	1.0		CA 180-405
Below Potrero Road	none						CA 05-175
Miscellaneous LA tributaries	none						all ML
Malibu Creek	none						all MA

(Continued)

Watershed/name of reach	TDS	Sulfate	Chloride	Boron	Nitrogen	SAR	Stations
Ballona Creek	none						all BA
Dominguez Channel	none						all DO
Los Angeles River Watershed:							
Above Figueroa Street	950	300	150		8		LA 207-475 (except 297-308, 325-337, 310-317)
Between Figueroa Street and Willow Street	1500	350	150		8		LA 05-206 (except 40-118, 200-205)
Below Willow Street	none						LA 03
Rio Hondo at Santa Ana Fwy.	750	300	150		8		LA 40-118 (except 105-112, 82-84)
Arroyo Seco above spreading grounds	300	40	15				LA 202-205
Eaton Canyon above Eaton Dam	250	30	10				LA 82-84
Big Tujunga above Hanson Dam	350	50	20				LA 325-337
Santa Anita Creek above Santa Anita spreading grounds	250	30	10				LA 105-112
Pacoima Wash above Pacoima spreading grounds	250	30	10				LA 310-317
All other minor San Gabriel Mountain Streams	300	40	15				LA 297-308
San Gabriel River Watershed:							
Above Morris Dam	250	30	10	0.6	2	2	SG 210-242
Between Morris Dam and Firestone Boulevard	750	300	150		8		SG 80-207 (except 203)
Firestone Boulevard to Willow	none						SG 66-75
Below Willow	none						SG 02-65

Note: Watershed includes all tributaries, so each reach includes all tributaries.

Excursion Analysis

The goal of the excursion analysis is to determine the number of times a water quality objective was exceeded per year, for each station and constituent. Stations were located into specific stream reaches as defined by the existing Basin Plan. Reaches and objectives were provided by the RWQCB, Los Angeles Region (see Tables 4 and 5).

The objective for pH is the range 6.5 to 8.5, and the objective for Nitrate is 45 mg/l as NO_3. It should be noted that the Sodium Absorption Ratio (SAR) and temperature were not included in the excursion analysis due to lack of uniform data availability.

Table 5. Dissolved Oxygen Criteria

RIVER/STREAM	STATIONS	COLD	SPAWN
Ventura River	VE 05-31, 36-53, 68-90	X	X
San Antonio Creek	VE 55-67	X	X
San Paula Creek	SC 180-205, 220	X	X
Sisar Creek	SC 210-215	X	X
Sespe Creek	SC 280-350	X	X
Hopper Creek	SC 395-430	X	
Piru Creek	SC 480-580, 605-615, 620	X	X
Canada de los Alamos	SC 583, 585	X	
Lockwood Creek	SC 625-665	X	
Bouquet Creek	NO STATIONS	X	
Malibu Creek	MA 05-10	X	
Pacoima Creek	LA 295-317	X	X
Big Tujunga Creek	LA 287-290, 320-337 except 325-328 & 330-333	X	X
Mill Creek	NO STATIONS	X	
Arroyo Seco	LA 202-205	X	
Big Santa Anita Creek	LA 110-112	X	X
Upper san Gabriel River	SG 205-210	X	
R. Fork San Gabriel River	SG 215-218, 225-227	X	X
N. Fork San Gabriel River	SG 230-235	X	
W. Fork San Gabriel River	SG 228, 230-242	X	X
Cattle Creek	SG 220	X	X
Coldwater Canyon Creek	NO STATIONS	X	X
Cow Creek	NO STATIONS	X	X
Devil's Canyon Creek	NO STATIONS	X	X
San Antonia Canyon Creek (LA Basin)	NO STATIONS	X	

Note: If a river/stream is designated as COLD, it has to meet 6 mg/L DO. If a river/stream is designed as both COLD and SPAWN, it must meet 7 mg/L DO. All other waters must meet 5 mg/L.

Outlier Evaluation

The possibility of excluding outlier data was considered in this study. Using the Los Angeles River as a test case, the number of sample points that exceeded various standard deviation offsets from the mean were determined, and listed in Table 6. Due to the unknown probability distributions of each constituent at each sampling location, all the data were retained within the overall data set, and the statistics were generated including all the sampling values.

Table 6. Number of Sampling Values Outside of Various Standard Deviation Offsets from the Mean (Los Angeles River Data)

Constituent	Number of Data Points	2 Standard Deviations	3 Standard Deviations	4 Standard Deviations	5 Standard Deviations
pH-Lab	11725	516	108	43	30
DO	7175	292	59	14	10
TDS	9694	457	95	42	29
SO4	8827	315	66	39	27
Chloride	12820	428	101	63	21
Boron	3280	150	78	36	21
Nitrate	8357	310	115	51	29
Nitrate-N	4464	149	53	12	6
Nitrite-N	1898	84	48	27	16
Phos-P	4060	159	79	35	27
pH-field	2546	108	24	0	0

Model Output

The output from the data base management model has been described in the previous section. The output formats for the statistical analysis and excursion analysis programs are described in the following sections.

Summary Output for the Statistical Analysis Program

The output for the statistical analysis program shows the station identification in the first row. Next, the basic statistical information (see Table 7) is provided in tabular form for each constituent. The basic statistical information includes: number of records, minimum sample value, 25th percentile, median, 75th percentile, 90th percentile, 99th percentile, maximum sample value, mean, and standard deviation. This information is provided for pre-1975, 1976-1985, and 1986-1993 time periods as well as for the entire data set. It should be noted that only the record number, minimum sample value, median, maximum sample value, and mean are provided for samples where the number of data are less than or equal to 10.

Table 7. Example of Statistical Program Output

**Station Name: LA003

Period	Constituent	Record	Min.	(25p)th	Median	(75p)th	(90p)th	(95p)th	(99p)th	Max.	Mean	St.Dev.
pre-1975	ph-Lab	268	4.30	7.40	7.70	8.00	8.40	8.60	10.19	12.70	7.74	.64
1976-1985	ph-Lab	26	7.10	7.67	8.10	8.42	8.69	9.03	9.10	9.10	8.09	.50
1986-1993	ph-Lab	0	←=== no statistics being calculated.									
all data	pH-Lab	294	4.30	7.40	7.70	8.00	8.40	8.60	9.36	12.70	7.78	.64
* Basin Plan Objective : 6.5 – 8.5												
pre-1975	DO	244	.10	1.40	3.00	6.57	9.15	11.88	15.68	18.60	4.16	3.65
1976-1985	DO	57	.50	3.30	5.60	7.60	9.04	9.22	10.40	10.40	5.63	2.53
1986-1993	DO	0	←=== no statistics being calculated.									
all data	DO	301	.10	1.60	3.60	6.70	9.08	10.69	14.79	18.60	4.44	3.51
* Basin Plan Objective : NOT DEFINED												
pre-1975	TDS	133	168.00	1100.00	3489.00	7918.50	9357.00	9714.30	9954.32	9972.00	4408.89	3368.93
1976-1985	TDS	38	356.00	767.25	852.00	1351.50	7841.10	9184.25	9512.00	9512.00	1897.11	2562.79
1986-1993	TDS	0	←=== no statistics being calculated.									
all data	TDS	171	168.00	969.00	1830.00	7299.00	9227.80	9562.00	9934.56	9972.00	3850.72	3367.35
* Basin Plan Objective : NOT DEFINED												
pre-1975	SO4	263	5.00	190.00	328.00	907.00	1408.00	1697.00	2170.56	5656.00	595.54	611.12
1976-1985	SO4	56	121.00	213.25	283.50	1022.25	1881.80	2089.55	2326.00	2326.00	679.93	639.35
1986-1993	SO4	0	←=== no statistics being calculated.									
all data	SO4	319	5.00	195.00	316.00	932.00	1527.00	1774.00	2190.00	5656.00	610.36	615.98
* Basin Plan Objective : NOT DEFINED												
pre-1975	Chloride	248	23.00	655.00	4053.00	6565.00	8224.40	8827.05	9803.06	9844.00	3972.25	3028.94
[illegible]-1985	Chloride	46	7.60	126.25	193.00	2435.75	7301.00	9074.60	9881.00	9881.00	1775.60	2874.48
[illegible]6-1993	Chloride	0	←=== no statistics being calculated.									
all data	Chloride	294	7.60	360.00	3705.00	6204.50	8167.50	8946.50	9807.90	9881.00	3628.55	3105.18
* Basin Plan Objective : NOT DEFINED												
pre-1975	Boron	41	.19	.64	3.60	13.50	19.28	39.00	40.00	40.00	8.16	10.09
1976-1985	Boron	46	←=== no statistics being calculated.									
1986-1993	Boron	0	←=== no statistics being calculated.									
all data	Boron	41	.19	.64	3.60	13.50	19.28	39.00	40.00	40.00	8.16	10.09
* Basin Plan Objective : NOT DEFINED												
pre-1975	Nitrate	187	.10	1.10	4.80	12.60	19.60	43.62	143.86	172.20	11.34	23.52
1976-1985	Nitrate	55	.20	2.60	10.10	17.50	22.20	23.86	27.00	27.00	10.97	7.80
1986-1993	Nitrate	0	←=== no statistics being calculated.									
all data	Nitrate	242	.10	1.30	5.80	14.00	20.28	35.40	139.14	172.20	11.25	20.99
* Basin Plan Objective : 45.00												
pre-1975	pH (field)	53	7.20	7.55	7.80	8.30	8.40	8.50	8.50	8.50	7.86	.40
1976-1985	pH (field)	31	7.30	7.80	8.00	8.20	8.82	8.90	8.90	8.90	8.05	.38
1986-1993	pH (field)	0	←=== no statistics being calculated.									
all data	pH (field)	84	7.20	7.60	7.90	8.20	8.45	8.50	8.90	8.90	7.93	.40
* Basin Plan Objective : 6.5 – 8.5												

Summary Output for the Excursion Analysis Program

The output (see Table 9) for the excursion analysis program includes the station identification. Next, the objectives for the constituents are listed. Finally, the number of exceedances and the number of observations for each constituent, for each year, are provided in tabular form.

Table 8. Example of Excursion Analysis Program Output

** Excursion Analysis for Station: LA003

	pH-Lab	DO	TDS	SO4	Chloride	Boron	Nitrate	Nitrogen	pH(field)
Objective	6-5.8.5	.0	.0	.0	.0	.0	45.0	.0	6.5-8-5
Year									
1951	0(9)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
1952	4(16)	0(0)	0(0)	0(0)	0(0)	0(0)	0(3)	0(0)	0(0)
1953	5(14)	0(0)	0(0)	0(0)	0(0)	0(0)	0(1)	0(0)	0(0)
1954	2(9)	0(0)	0(0)	0(0)	0(0)	0(0)	0(3)	0(0)	0(1)
1955	0(7)	0(0)	0(0)	0(0)	0(0)	0(0)	0(2)	0(0)	0(2)
1956	0(14)	0(0)	0(0)	0(0)	0(0)	0(0)	0(3)	0(0)	0(3)
1957	0(15)	0(0)	0(0)	0(0)	0(0)	0(0)	0(2)	0(0)	0(2)
1958	1(17)	0(0)	0(0)	0(0)	0(0)	0(0)	1(10)	0(0)	0(2)
1959	0(12)	0(0)	0(0)	0(0)	0(0)	0(0)	1(13)	0(0)	0(2)
1960	1(11)	0(0)	0(0)	0(0)	0(0)	0(0)	0(7)	0(0)	0(2)
1961	1(13)	0(0)	0(0)	0(0)	0(0)	0(0)	2(9)	0(0)	0(1)
1962	0(11)	0(0)	0(0)	0(0)	0(0)	0(0)	0(5)	0(0)	0(2)
1963	0(10)	0(0)	0(0)	0(0)	0(0)	0(0)	1(7)	0(0)	0(5)
1964	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(8)	0(0)	0(11)
1965	0(4)	0(0)	0(0)	0(0)	0(0)	0(0)	0(8)	0(0)	0(8)
1966	1(9)	0(0)	0(0)	0(0)	0(0)	0(0)	0(10)	0(0)	0(4)
1967	0(11)	0(0)	0(0)	0(0)	0(0)	0(0)	2(12)	0(0)	0(2)
1968	0(13)	0(0)	0(0)	0(0)	0(0)	0(0)	2(13)	0(0)	0(1)
1969	0(10)	0(0)	0(0)	0(0)	0(0)	0(0)	0(12)	0(0)	0(2)
1970	0(10)	0(0)	0(0)	0(0)	0(0)	0(0)	0(10)	0(0)	0(0)
1971	0(13)	0(0)	0(0)	0(0)	0(0)	0(0)	0(11)	0(0)	0(1)
1972	0(12)	0(0)	0(0)	0(0)	0(0)	0(0)	0(8)	0(0)	0(2)
1973	0(10)	0(0)	0(0)	0(0)	0(0)	0(0)	0(8)	0(0)	0(0)
1974	2(12)	0(0)	0(0)	0(0)	0(0)	0(0)	0(11)	0(0)	0(0)
1975	0(11)	0(0)	0(0)	0(0)	0(0)	0(0)	0(11)	0(0)	0(0)
1976	1(6)	0(0)	0(0)	0(0)	0(0)	0(0)	0(6)	0(0)	0(0)
1977	0(2)	0(0)	0(0)	0(0)	0(0)	0(0)	0(7)	0(0)	0(6)
1978	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(11)	0(0)	2(11)
1979	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(12)	0(0)	1(12)
1980	1(7)	0(0)	0(0)	0(0)	0(0)	0(0)	0(9)	0(0)	0(2)
1981	2(11)	0(0)	0(0)	0(0)	0(0)	0(0)	0(10)	0(0)	0(0)

Notes:
1. A zero in the objective indicates either the objective value is not defined of the data set is empty.
2. First value in the excursion table indicates number of exceedances of objectives per year.
3. The value in the parentheses indicates number of observations per year.

Graphical Displays

Box-and-Whisker Plots

A box-and-whisker plot (see Figure 1) shows five values relevant to a data set, such as the median, the maximum, the minimum, and inter quartile values (25^{th} and 75^{th} percentiles). The 25^{th} and 75^{th} percentiles form the limits of the box. The median value is shown inside that box. The lower whisker extends to the minimum value of the sample set and the upper whisker extends to the maximum value of the sample set. It is noted that generally whiskers extend to 1.5 times the interquantile range (IQR); however in the subject output, the entire sample data set is used.

Trend Analysis Plots

A trend refers to an increase or decrease in sample values of a constituent over time. A dashed line on the trend analysis plot (see Figure 2) represents the trend of a data set. The rate of increase or decrease in sample values over time can be represented by the slope of the dashed line which is computed using a standard least squares analysis.

Discussion

A computer program was developed to manage a massive surface water-quality monitoring database and to provide basic statistical information and excursion analysis results for the California Regional Water Quality Control, Los Angeles Region. These analyses provided the Regional Board several useful management tools for use in developing new water quality standards, assessing water quality trends, and analyzing exceedences of water quality objectives.

The water quality monitoring data provided from the data management program could be easily imported into a statistical analysis program for graphical display and other statistical analyses. The compilation of the total available water quality data set for some 1500 sample sites provides a tremendous source of data for further research into water quality issues as they pertain to hydrological processes. These data are available from the RWQCB, Los Angeles Region and are suitable for additional statistical analysis.

References

Water Quality Control Plan, Los Angeles Region, Abstracts and Appendices of 1975 Plans (printed with 1978, 1990, 1991 and 1992 changes), California Regional Water Quality Control Board, Los Angeles Region. 1992.

Water Quality Control Plan, Los Angeles Region, Basin Plan for the Coastal Watersheds of Los Angeles and Ventura Counties. 1994. California Regional Water Quality Control Board, Los Angeles Region.

Hogg, R.V. and Tanis, E.A. 1997. "Probability and Statistical Inference," Macmillan Publishing Company, Ltd.

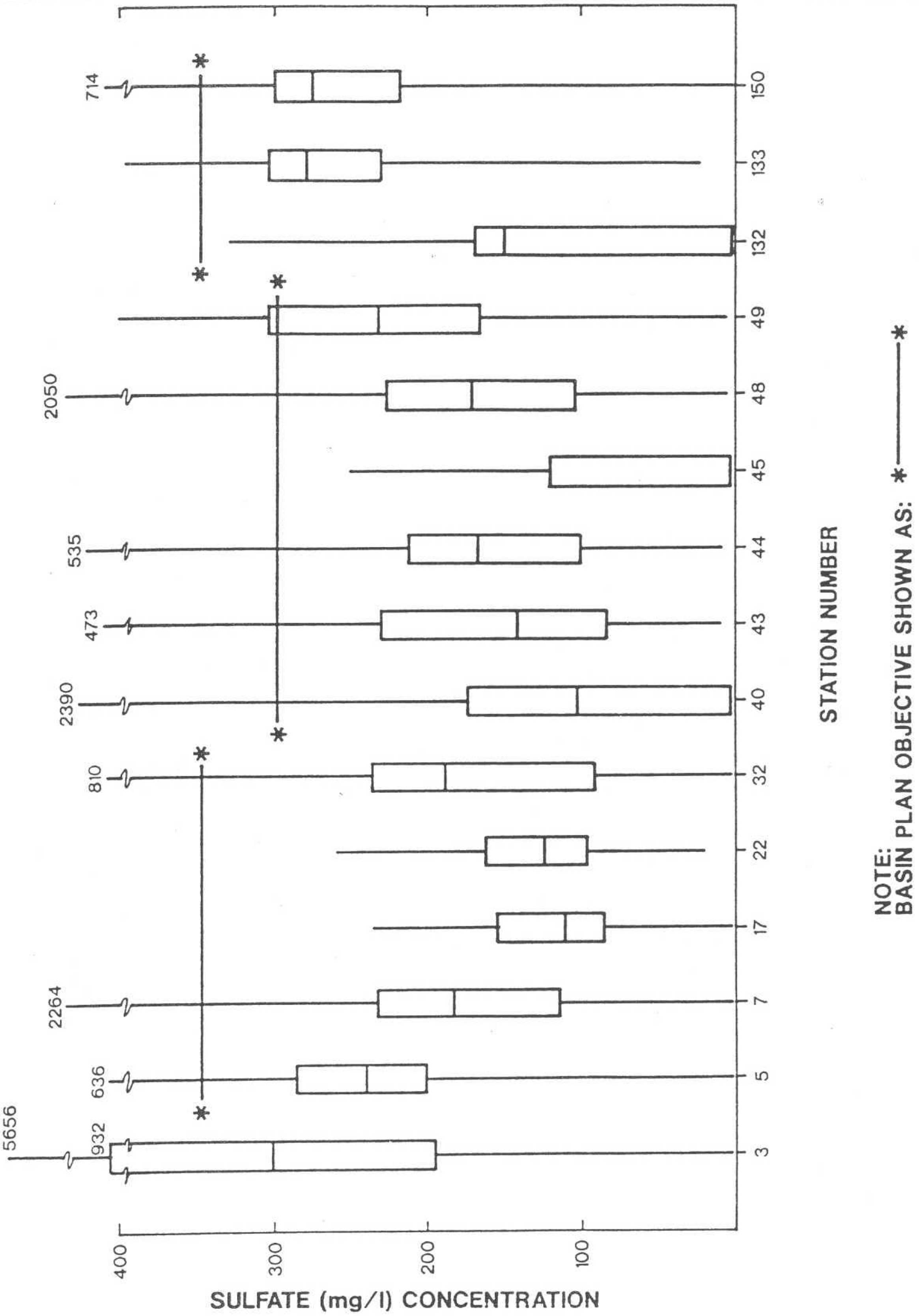

Figure 1. Box and Whisker Plots of Sulfate Concentration for Los Angeles River

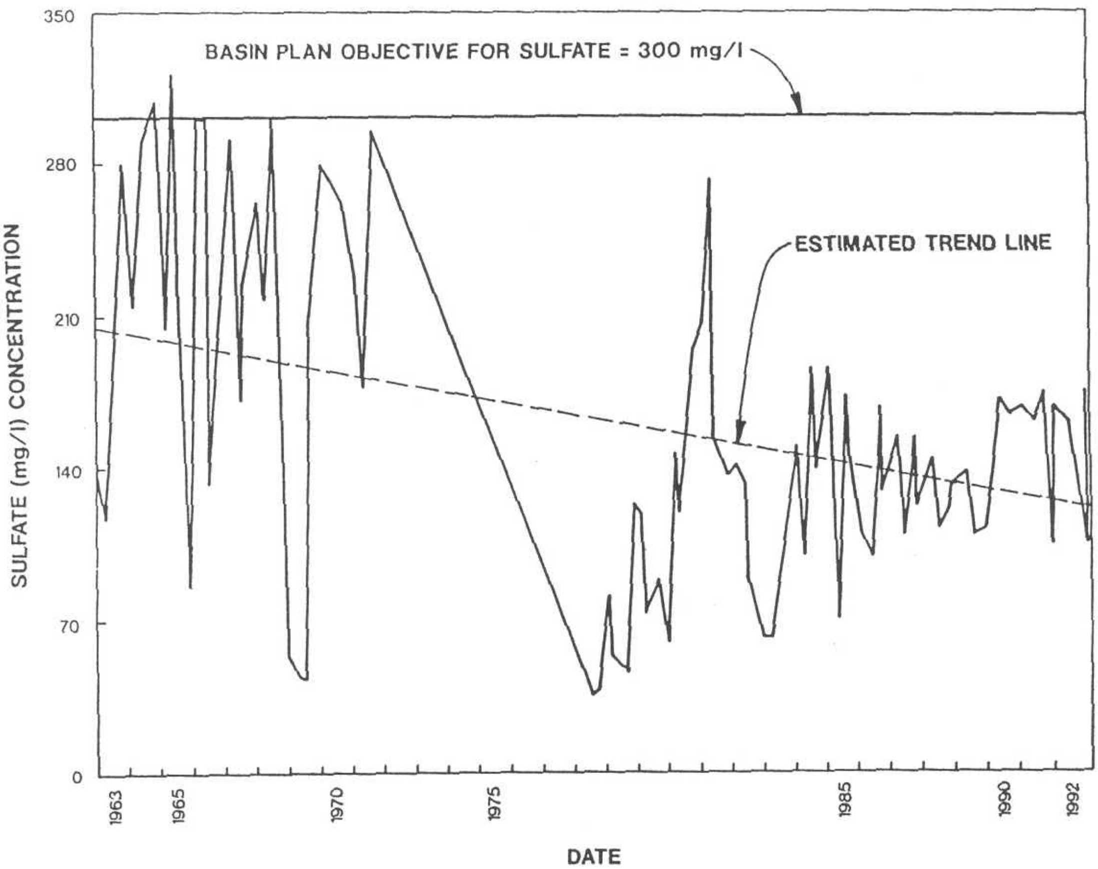

Figure 2. Trend Analysis Plot of Sulfate Concentration