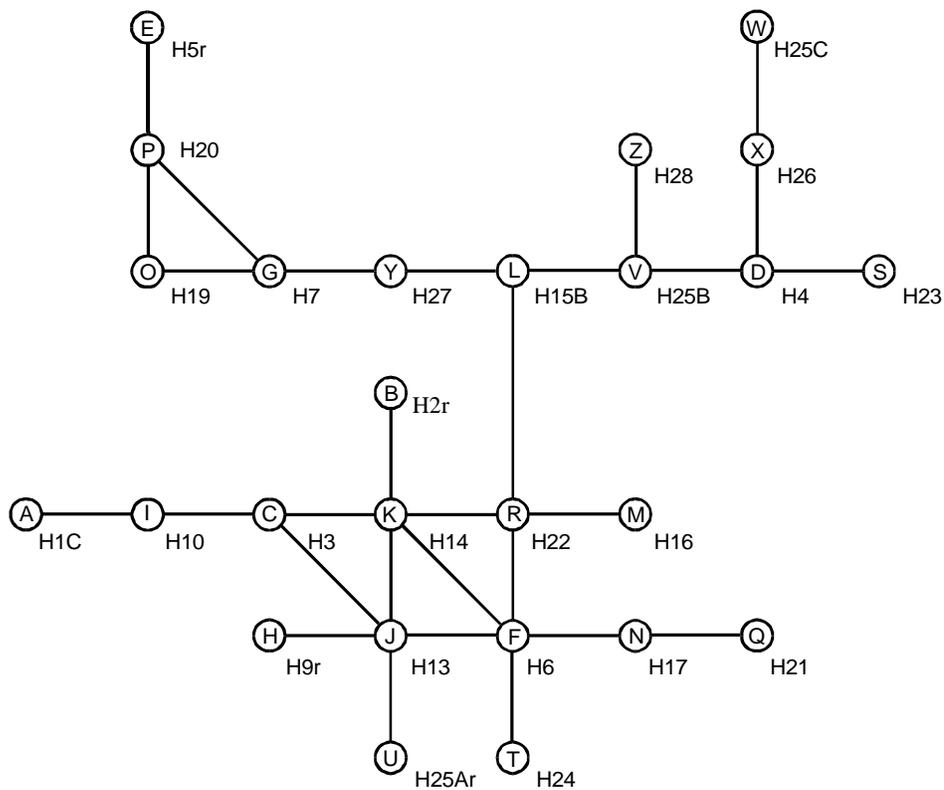


A Temporal and Spatial Analysis of Ground-Water Levels for Effective Monitoring in Huron County, Michigan

U.S. GEOLOGICAL SURVEY
Open-File Report 98-629



Prepared in cooperation with
Huron County, Michigan



A Temporal and Spatial Analysis of Ground-Water Levels for Effective Monitoring in Huron County, Michigan

By D.J. Holtschlag and M.J. Sweat

**U.S. GEOLOGICAL SURVEY
Open-File Report 98-629**

**Prepared in cooperation with
Huron County, Michigan**

**Lansing, Michigan
1999**



U.S. DEPARTMENT OF THE INTERIOR

BRUCE BABBITT, *Secretary*

U.S. GEOLOGICAL SURVEY

Charles G. Groat, *Director*

For additional information
write to:

District Chief
U.S. Geological Survey, WRD
6520 Mercantile Way, Suite 5
Lansing, MI 48911-5991

Copies of this report can be
purchased from:

U.S. Geological Survey
Branch of Information Services
Box 25286
Denver, CO 80225-0286

Contents

	Page
Abstract.....	1
Introduction.....	1
Description of the Study Area and Monitoring Wells.....	1
Acknowledgments	3
Temporal Components of Ground-Water Level Measurements	3
Structural Time-Series Analysis.....	3
Results of Time-Series Analysis.....	6
Spatial Interrelations Among Ground-Water Level Measurements	15
A Graphical Model of the Covariance Structure	16
Results of Graphical Modeling.....	18
Achieving Effectiveness by Ranking Monitoring Wells, and Specifying Network	
Constraints	18
Summary.....	22
References Cited.....	24

Figures

1. Map showing the location of monitoring wells in Huron County, Michigan	2
2. Diagram showing the stratigraphic succession and aquifer nomenclature in Huron County, Michigan	4
3-8. Graphs showing water levels and time-series components at wells:	
3. H1C, H2r, H3, H4, and H5r.....	9
4. H6, H7, H9r, H10, and H13.....	10
5. H14, H15B, H16, H17, and H19	11
6. H20, H21, H22, H23, and H24	12
7. H25Ar, H25B, H25C, and H26	13
8. H27 and H28.....	14
9. Signs of the elements of the empirical covariance matrix.....	17
10. Graph showing the relation between the number of edges in the graphical model and the Bayesian Information Criteria	19
11. Diagram of the independence graph of water-level innovations at monitoring wells.....	20
12. Map showing the geographic reference for the independence graph in Huron County	21
13. Graph showing the relation between the number of wells monitored and network uncertainty	23

Tables

1. Selected attributes of monitoring wells in Huron County, Michigan	5
2. Structural time-series components of water levels in wells in Huron County, Michigan.....	7

CONVERSION FACTORS AND VERTICAL DATUM

Multiply	By	To Obtain
inch (in)	25.4	millimeter
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
square mile (mi ²)	2.590	square kilometer

Temperature in degrees Fahrenheit (°F) can be converted to degrees Celsius (°C)
as follows: $^{\circ}\text{C} = 5/9 (^{\circ}\text{F} - 32)$.

Vertical datum: In this report “sea level” refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

A Temporal and Spatial Analysis of Ground-Water Levels for Effective Monitoring in Huron County, Michigan

By D.J. Holtschlag *and* M.J. Sweat

ABSTRACT

Quarterly water-level measurements were analyzed to assess the effectiveness of a monitoring network of 26 wells in Huron County, Michigan. Trends were identified as constant levels and autoregressive components were computed at all wells on the basis of data collected from 1993 to 1997, using structural time series analysis. Fixed seasonal components were identified at 22 wells and outliers were identified at 23 wells. The 95-percent confidence intervals were forecast for water-levels during the first and second quarters of 1998. Intervals in the first quarter were consistent with 92.3 percent of the measured values. In the second quarter, measured values were within the forecast intervals only 65.4 percent of the time. Unusually low precipitation during the second quarter is thought to have contributed to the reduced reliability of the second-quarter forecasts.

Spatial interrelations among wells were investigated on the basis of the autoregressive components, which were filtered to create a set of innovation sequences that were temporally uncorrelated. The empirical covariance among the innovation sequences indicated both positive and negative spatial interrelations. The negative covariance components are considered to be physically implausible and to have resulted from random sampling error. Graphical modeling, a form of multivariate analysis, was used to model the covariance structure. Results indicate that only 29 of the 325 possible partial correlations among the water-level innovations were statistically significant. The model covariance matrix, corresponding to the model partial correlation structure, contained only positive elements. This model covariance was sequentially partitioned to compute a set of partial covariance matrices that were used to rank the effectiveness of the 26 monitoring wells from greatest to least. Results, for example, indicate that about 50 percent of the uncertainty of the water-

level innovations currently monitored by the 26-well network could be described by the 6 most effective wells.

INTRODUCTION

A network of 26 ground-water-level monitoring wells in Huron County, Michigan is operated by the County in cooperation with the U.S. Geological Survey. Information from the network is used to assess long-term and seasonal variations in ground-water levels and local changes that may be partly associated with human activities. Results from monitoring provide county planners and water-resources managers with data needed to develop ground-water-management plans that are consistent with available resources and expected demands.

This report describes a technique for assessing and improving the effectiveness of ground-water-level monitoring networks. The technique provides a statistical basis for extracting information on trends and seasonal variations in water levels at individual wells, forecasting ground-water levels, and assessing the potential redundancy in measurements of spatially-correlated phenomena. Also, the technique provides a basis for improving the effectiveness of data collection by identifying specific subsets of monitoring wells that provide data which is consistent with informational needs and available monitoring resources. The technique was applied to a network of 26 monitoring wells in Huron County, Michigan on the basis of quarterly water-level monitoring from 1993 to 1998.

Description of the Study Area and Monitoring Wells

Huron County is in the east-central part of Michigan's Lower Peninsula (fig. 1). The county, which is shaped roughly like a semicircle, is bounded along the north by 91 mi of Great Lake shoreline (Lake Huron and Saginaw Bay) and along the south by Sanilac and Tuscola Counties. Most streams start within the county and flow to the lake or bay. Land surface is flat to rolling; elevations range from 580 ft above sea level along the shoreline to more than 800 ft near Ubyly, Michigan. The county has an area of 830 mi², most of which is pasture and cropland; population in 1990 was 34,951.

The geology of Huron County consists of consolidated strata of Mississippian and Pennsylvanian age and unconsolidated surficial deposits of Pleistocene age. From oldest to youngest, the consolidated strata are the Coldwater Shale, the Marshall Formation, the Michigan Formation, the Bayport Limestone, and the Saginaw Formation (fig. 2). Throughout the county, unconsolidated lake and glacial deposits overlie the consolidated deposits with maximum thicknesses that range between 80 and 120 ft.

In 1990, the U.S. Geological Survey (USGS) completed a study of the hydrogeology of Huron County, Michigan (Sweat, 1991). In 1993, Huron County and the USGS agreed to continue water-level monitoring at 26 selected wells (fig. 1) throughout Huron County (Sweat, 1995, 1996, 1997). The wells monitor hydraulic head (water levels) in four geologic units: 3 wells measure heads in the Coldwater confining unit, 17 wells measure heads in the Marshall aquifer, 5 wells measure heads in the Saginaw Formation, and 1 well measures head in the glaciofluvial aquifer (table 1). Of the 26 wells operated, water levels are recorded continuously in 4 wells and measured quarterly in 22 wells. First quarter (January through March) measurements were generally made in March; second-, third-, and fourth-quarter measurements were generally made in June, September, and December, respectively. Measured depths to water below measuring points were converted to elevations above sea level by subtracting depths from the elevation of the measuring point.

Acknowledgments

James LeCureux and Carol Schadd of Huron County Cooperative Extension Service assisted the project by obtaining many of the field measurements of water level. Fred Nurnberger, State Climatologist of Michigan, provided precipitation data for Huron County, Michigan.

TEMPORAL COMPONENTS OF GROUND-WATER-LEVEL MEASUREMENTS

Repetitive ground-water level measurements track persistent and ephemeral changes in water levels through time. Persistent changes may include trend and seasonal components that characterize the ground-water resource and provide information needed for assessment and effective

management. Ephemeral changes reflect day to day changes that are influenced by recent weather conditions, such as the amount of time since the last soaking rain. Ephemeral changes obscure detection of persistent patterns.

Time-series analysis provides a systematic and consistent basis for identifying trends and seasonal components of water levels, quantifying autoregressive (serial correlation) characteristics, and describing the variability associated with ephemeral changes. Classical time-series analysis (Brockwell and Davis, 1987) describe time series by use of ARIMA (autoregressive integrated moving average) models, following removal of trend components. ARIMA models, however, are based on statistical assumptions that are often difficult to verify for hydrologic time series and generally require at least 50 and preferably 100 observations for estimation (Box and Jenkins, 1976, p. 18). Structural time-series analysis provide an alternative that can simultaneously identify trends, seasonal components, autoregressive characteristics, and the influences of explanatory variables on time series with fewer observations than ARIMA models. Moreover, the description of trend and seasonal components is flexible (not limited to simplified deterministic forms) and is directly interpretable from a physical perspective.

Structural Time-Series Analysis

Structural time-series analysis (Harvey, 1994) as implemented by Koopman and others (1995) was used to identify trend, seasonal, and autoregressive components in quarterly water-level measurements made in Huron County between 1993 and 1997; outliers (unusual values) also were identified. Resulting models were used to forecast the expected values and uncertainties of water levels for the six quarters following the data used for model development. Forecast intervals were compared with data obtained during the first two quarters of 1998 to assess model adequacy.

The general form of the structural time series model is:

$$y_t = \mu_t + \gamma_t + \sum_{i \in I} \lambda_i w_{i,t} + v_t \quad (1)$$

where

y_t is the water-level measurement at time t , and μ_t is the trend component at time t . In structural

ERATHEM	SYSTEM	SERIES	GROUP	STRATIGRAPHIC UNIT	HYDROGEOLOGIC UNIT	THICKNESS (feet)	
CENOZOIC	Quaternary	Pleistocene		Glacial and lacustrine deposits	Surficial deposits (aquifer)	0-10	
				Glacial deposits		0-120	
				Unconformity			
PALEOZOIC	Pennsylvanian	Middle		Saginaw Formation	Aquifer system	0-100	
		Lower					
	Unconformity						
	Mississippian	Meramecian	Grand Rapids	Bayport Limestone	Confining units	0-100	
				Michigan Formation		0-175	
		Osagean		Marshall Formation	Napoleon Sandstone Member of Marshall Formation (aquifer)	0-120	
					Sandstones in lower part of Marshall Formation (aquifer)	0-225	
Kinderhookian		Coldwater Shale	Confining unit	1,000-1,200			

Figure 2. Stratigraphic succession and aquifer nomenclature in Huron County, Michigan (Sweat, 1991).

time-series models, trend is described by level (μ_t) and slope (β_t) components, which may be either stochastic or deterministic. When the level is stochastic, an uncorrelated random sequence η_t with a mean of zero and variance σ^2_{η} , is introduced. Similarly, when the slope is stochastic, an uncorrelated random sequence ζ_t with mean zero and variance σ^2_{ζ} is present. (The random sequences η_t and ζ_t also are assumed to be uncorrelated with each other). The effect of η_t is to allow the level to shift up and down locally, while ζ_t allows the slope to change. In the special case where either σ^2_{η} or σ^2_{ζ} are equal to zero, the corresponding components are considered deterministic and the associated trend component is uniform throughout the series. In the special case where the slope parameter, β , and σ^2_{η} are equal to zero, the trend is a constant (level). The state-space representation of the trend component used in the structural model is:

$$\begin{bmatrix} \mu_{i,t} \\ \beta_{i,t} \end{bmatrix} = \begin{bmatrix} \mu_{i,t-1} + \beta_{i,t-1} + \eta_{i,t} \\ \beta_{i,t-1} + \zeta_{i,t} \end{bmatrix}. \quad (2)$$

γ_t is the seasonal component in the q^{th} quarter at time t . As in the case of the trend component, the structural time series model provides the flexibility for stochastic seasonality. The general form of the seasonal component is:

$$\gamma_t = \sum_{q=1}^4 \delta(Q(t) - q) \cdot \gamma_q + \omega_t. \quad (3)$$

γ_1, γ_2 , and γ_3 are estimated seasonal coefficients for the first, second, and third quarters, respectively, corresponding to the periods January through March, April through June, and July through September. Because the seasonal components were restricted to sum to zero, the fourth seasonal component, γ_4 (October through December) was set equal to $-(\gamma_1 + \gamma_2 + \gamma_3)$. Here, the function $Q(t)$ returns the quarter corresponding to time t , and $\delta(\cdot)$ is the Kronecker delta function which returns a 1 if $(Q(t) - q)$ evaluates to zero, and is zero otherwise. Possible stochastic disturbances in seasonality are described by the disturbance ω_t , which is an uncorrelated random sequence with a mean of zero and variance of σ^2_{ω} . When σ^2_{ω} is equal to zero, the

Table 1. Selected attributes of monitoring wells in Huron County, Michigan

[G, indicates well in glaciofluvial aquifer; S, indicates well in the Saginaw aquifer system;

M, indicates well in Marshall aquifer; and C, indicates Coldwater confining unit.

Well numbers ending in an "r" indicate that water levels are recorded continuously.]

Vertex	Well number	Latitude	Longitude	Altitude (feet above sea level)	Hole depth (feet below land surface)	Aquifer code
A	H1C	43°41'44"	83°27'44"	600.00	76	S
B	H2r	43°41'03"	83°13'03"	747.60	91	G
C	H3	43°42'19"	83°10'42"	731.70	120	M
D	H4	43°44'02"	83°05'59"	751.60	80	M
E	H5r	43°43'23"	82°56'19"	796.03	171	M
F	H6	43°45'26"	82°50'26"	781.50	90	M
G	H7	43°42'08"	82°41'11"	726.80	140	C
H	H9r	43°49'47"	83°23'33"	584.20	180	S
I	H10	43°49'43"	83°19'08"	617.07	150	S
J	H13	43°49'00"	83°13'16"	642.35	120	S
K	H14	43°49'00"	83°07'18"	681.30	100	M
L	H15B	43°48'11"	83°02'16"	751.20	99	M
M	H16	43°48'53"	82°56'52"	771.50	160	M
N	H17	43°48'39"	82°49'24"	751.00	80	M
O	H19	43°52'42"	83°15'39"	611.90	100	M
P	H20	43°53'26"	83°10'57"	631.00	60	S
Q	H21	43°54'48"	83°01'27"	702.90	80	M
R	H22	43°52'13"	82°57'47"	695.50	150	M
S	H23	43°54'45"	82°48'42"	721.80	150	C
T	H24	43°54'11"	82°44'42"	691.50	100	C
U	H25Ar	43°57'36"	83°09'48"	600.80	200	M
V	H25B	43°57'36"	83°09'48"	601.00	160	M
W	H25C	43°57'36"	83°09'48"	602.20	40	M
X	H26	43°56'42"	83°02'55"	662.70	60	M
Y	H27	43°57'33"	82°54'58"	716.50	80	M
Z	H28	43°57'33"	82°51'22"	691.70	75	M

seasonal component is deterministic (fixed) and $\gamma_t = \gamma_q \cdot \sum_{i \in I} \lambda_i w_{i,t}$ is a term that accounts for individual outliers indexed by the elements i of the set of outliers $\{I\}$. The scalar λ_i coefficients indicate the magnitude of outliers corresponding to the pulse intervention vectors represented by $w_{i,t}$. The individual vectors take on a value of one during periods of intervention (corresponding to outliers) and are zero otherwise. Finally, v_t is the autoregressive error component of the form:

$$v_t = \phi v_{t-1} + \varepsilon_t, \quad (4)$$

where ϕ is the autoregressive coefficient for components that are separated by one season (lag 1). In the special case where ϕ is equal to 0, the error v_t is an uncorrelated disturbance sequence with mean zero and variance σ_ε^2 . In this analysis, ε_t was computed as $v_t - \phi v_{t-1}$, and is referred to subsequently as the innovation sequence.

In the analysis for each monitoring well, time-series models potentially included stochastic level, slope, and seasonal components and an autoregressive component. Model components were eliminated in a stepwise procedure to minimize the Bayesian Information Criteria (*BIC*), a statistic that decreases with model-error variance and increases with the number of (fixed) parameters and stochastic components (hyperparameters) in the model. In general, stochastic components were eliminated before deterministic components. Standardized residuals that had a magnitude greater than 2.0 were examined to identify possible outliers. To confirm the presence of outliers, models were computed with and without intervention variables corresponding to possible outliers. The model with the minimum *BIC* value was selected. In addition to estimating intervention effects, the structural model also has the flexibility to estimate the effects of explanatory (independent) variables on water levels.

Results of Time-Series Analysis

The identified time-series models were simple and remarkably consistent from well to well (table 2 and figs. 3-8). All models included autoregressive error components with positive (AR 1) coefficients. None of the selected models included stochastic components in levels, slopes, or season-

als. In addition, no deterministic slope components were significant, thus all trend components are described as constant levels. Model structures differed only by the presence or absence of a fixed seasonal component and the number of outliers detected.

Seasonal components were identified at 22 of the 26 wells analyzed. Water levels in 4 wells completed in the Marshall formation had no seasonal components; however, water levels in 13 other wells in the Marshall formation had seasonal components. Inspection of seasonal components indicate that 20 of the first quarter and 21 of the second quarter coefficients were positive, indicating higher than average water levels in the first two quarters of the year. In contrast, seasonal coefficients were negative in all 22 wells during the third quarter, indicating lower than average water levels for this period. Results in the fourth quarter were mixed, with 11 wells having positive and 11 wells having negative coefficients, which were computed as the negative sums of the estimated coefficients.

A total of 35 outliers were identified at 23 wells; 28 of these outliers were negative and 7 were positive. The most outliers at any one well was three. Most (32) of the outliers occurred in the third quarter, 21 of these occurred in the third quarter of 1995 alone. The total number of outliers identified represents a larger percentage (6.7 percent) of the measurements than expected for a sample from a population described by a single normal probability distribution. The large number of outliers in 1995 may be partly related to precipitation patterns. Review of precipitation records at Bad Axe, Michigan in central Huron County indicates that third quarter precipitation of 6.95 inches in 1995 was far below average for the four remaining years of 12.96 inches. Inclusion of quarterly precipitation as a covariate in the structural time-series analysis, however, did not significantly reduce model error.

Water-level forecasts and associated uncertainties were computed on the basis of the model developed with five years of data available through the fourth quarter of 1997. During the first quarter of 1998, measured water levels were within the 95 percent of the lead-1 (one quarter ahead of the last measurement used in model development) forecast intervals in 92.3 percent of the cases (24 of 26

Table 2. Structural time-series components of water levels in wells in Huron County, Michigan [RMSE, Root mean square error; SEP, Standard error of prediction; AR--Autoregressive; Yr-Q, two-digit year code and quarter (January-March equals 1)]

Well number	Level (RMSE) (feet above sea level)	Seasonal Component			AR Coef. (σ_ϵ)	Outlier			SEP
		Coef. γ_1 (RMSE)	Coef. γ_2 (RMSE)	Coef. γ_3 (RMSE)		Date (YR-Q)	Coef. λ	RMSE	
H1C	583.78 (3.42)	-0.356 (.150)	0.688 (.143)	-0.298 (.157)	0.988 (.560)	95-3	-2.00	0.445	0.486
H2r	715.44 (.126)	0.354 (.179)	0.586 (.177)	-0.767 (.192)	0.146 (.477)	95-3	-1.43	.527	.415
H3	702.17 (.191)	0.078 (.137)	0.364 (.133)	-0.348 (.162)	0.538 (.410)	93-3 95-3	-1.03 -1.56	.417 .417	.345
H4	738.45 (.216)	0.126 (.163)	0.279 (.158)	-.294 (.173)	0.512 (.492)	95-3	-1.31	.490	.428
H5r	783.75 (.328)	--	--	--	0.310 (1.03)	--	--	--	1.012
H6	767.37 (.194)	0.431 (.167)	0.399 (.163)	-0.924 (.199)	0.451 (.483)	95-3 96-3	-1.14 1.52	.509 .509	.406
H7	709.83 (.086)	0.129 (.061)	0.205 (.059)	-0.216 (.053)	0.583 (.170)	96-2 97-1	-1.10 -2.84	.165 .166	.143
H9r	579.41 (.673)	0.166 (.180)	0.647 (.173)	-0.340 (.189)	0.829 (.625)	95-3	-1.68	.538	.544
H10	592.63 (.103)	0.228 (.093)	0.368 (.083)	-0.410 (.091)	0.469 (.251)	94-1 95-3	-5.23 -.55	.255 .254	.211
H13	609.89 (.245)	0.721 (.279)	0.618 (.275)	-1.490 (.299)	0.293 (.779)	95-3	-2.40	.836	.678
H14	676.27 (.220)	0.073 (.221)	0.421 (.217)	-0.545 (.236)	0.366 (.633)	95-3	-1.65	.665	.551
H15B	734.27 (.599)	0.258 (.277)	0.204 (.267)	-0.728 (.292)	0.694 (.904)	95-3	-2.78	.831	.787
H16	742.08 (.342)	0.061 (.175)	0.551 (.169)	-0.433 (.185)	0.661 (.564)	95-3	-1.20	.526	.491
H17	743.42 (.278)	1.497 (.249)	1.513 (.243)	-3.848 (.346)	0.436 (.692)	94-3 95-3 96-3	4.52 -16.45 5.42	.777 .777 .777	.561
H19	607.66 (.120)	0.219 (.098)	-0.496 (.096)	-0.335 (.096)	0.469 (.296)	--	--	--	.266

Table 2. Structural time-series components of water levels in wells in Huron County, Michigan--Continued
 [RMSE, Root mean square error; SEP, Standard error of prediction; AR--Autoregressive;
 Yr-Q, two-digit year code and quarter (January-March equals 1)]

Well number	Level (RMSE) (feet above sea level)	Seasonal Component			AR Coef. (σ_ε)	Date (YR-Q)	Outlier		SEP
		Coef. γ_1 (RMSE)	Coef. γ_2 (RMSE)	Coef. γ_3 (RMSE)			Coef. λ	RMSE	
H20	617.07 (.152)	-0.150 (.107)	0.497 (.104)	-0.244 (.104)	0.541 (.330)	--	--	--	0.297
H21	692.67 (.092)	0.282 (.111)	0.189 (.109)	-0.739 (.133)	0.251 (.300)	95-3 96-3	-1.05 1.51	0.336 .336	.252
H22	681.64 (.214)	0.402 (.131)	0.358 (.127)	-0.903 (.155)	0.604 (.403)	95-3 96-3	-1.34 1.40	.398 .398	.339
H23	712.25 (.101)	0.045 (.028)	0.011 (.027)	-0.071 (.029)	0.822 (.097)	95-3	-0.20	.084	.085
H24	665.87 (.128)	0.250 (.063)	0.236 (.061)	-.450 (.066)	0.676 (.203)	96-3	0.40	.188	.177
H25Ar	594.77 (.226)	--	--	--	0.278 (.735)	95-3	-2.19	.717	.701
H25B	598.33 (.210)	0.120 (.154)	0.345 (.150)	-.587 (.164)	0.523 (.468)	95-3	-1.78	.464	.408
H25C	598.50 (.115)	--	--	--	0.124 (.438)	93-3 95-3	-2.06 -2.71	.443 .443	.405
H26	658.21 (.174)	--	--	--	0.278 (.559)	93-3 95-3	-2.88 -5.08	.545 .545	.518
H27	683.57 (.293)	0.362 (.140)	0.256 (.135)	-0.589 (.166)	0.688 (.448)	95-3 97-3	-1.11 -5.05	.426 .426	.376
H28	671.98 (.141)	0.758 (.157)	0.423 (.154)	-1.387 (.187)	0.303 (.431)	95-3 96-3	-1.40 2.88	.476 .476	.362

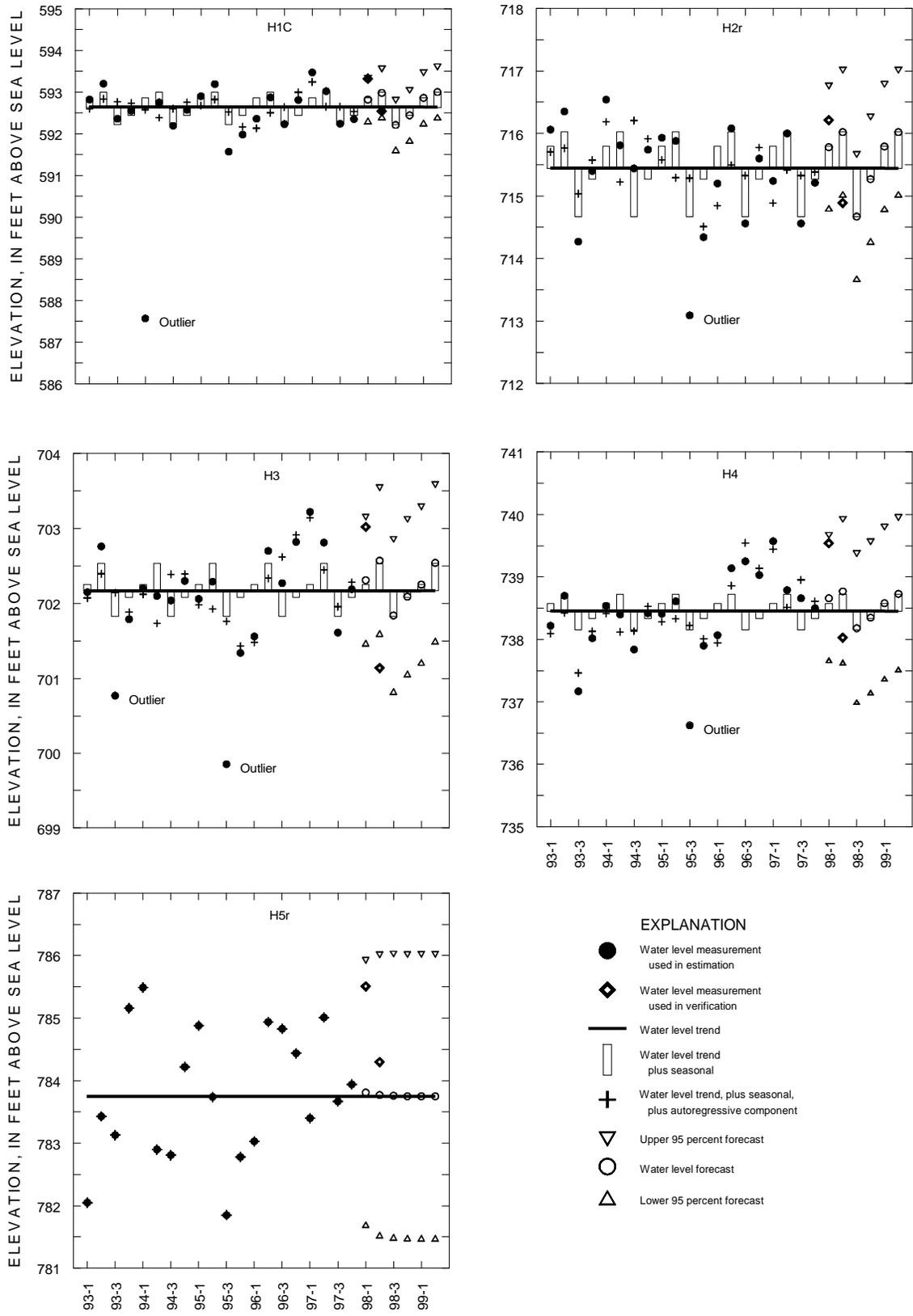


Figure 3. Water levels and time-series components at wells H1C, H2r, H3, H4, and H5r.

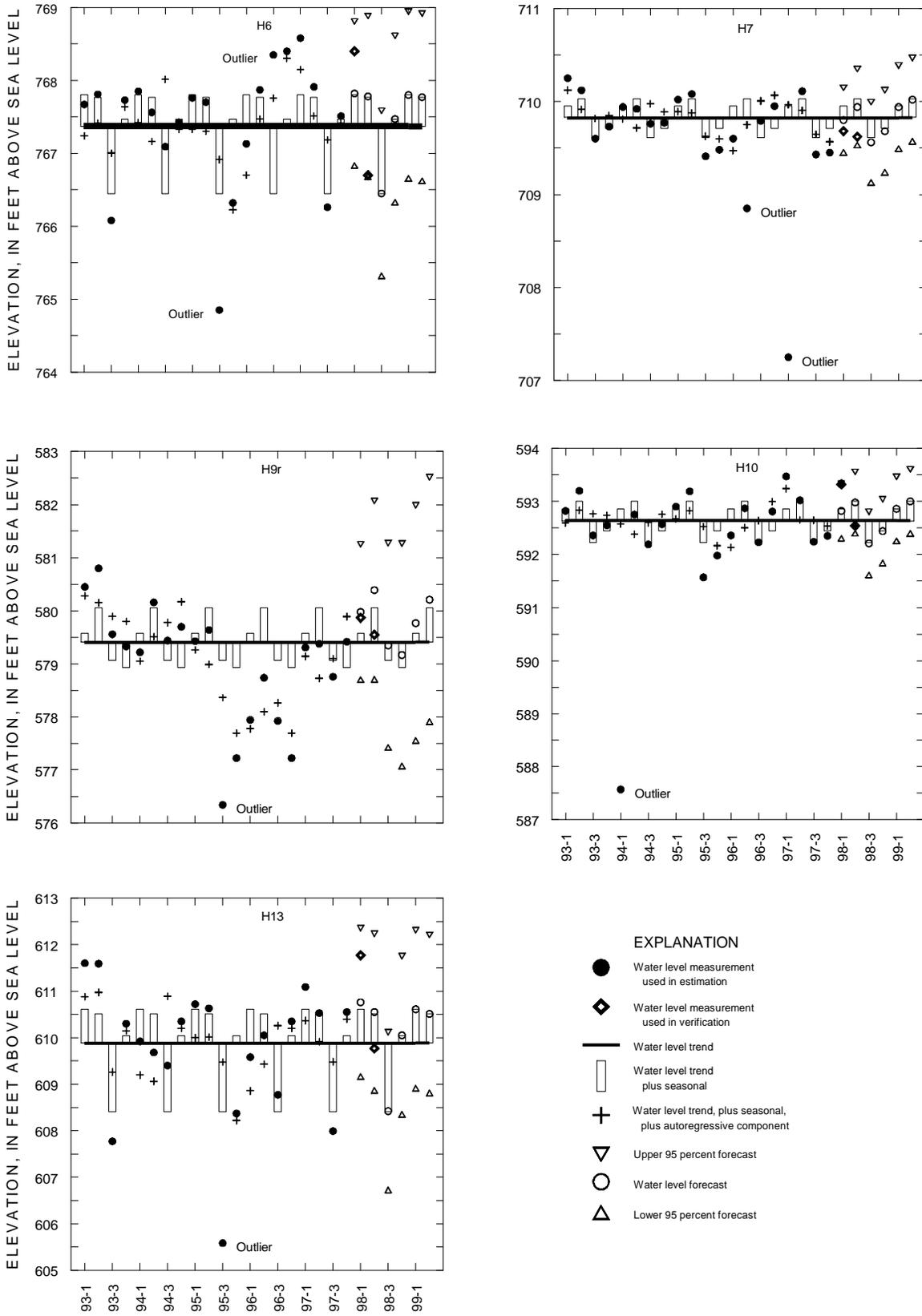


Figure 4. Water levels and time-series components at wells H6, H7, H9r, H10, and H13.

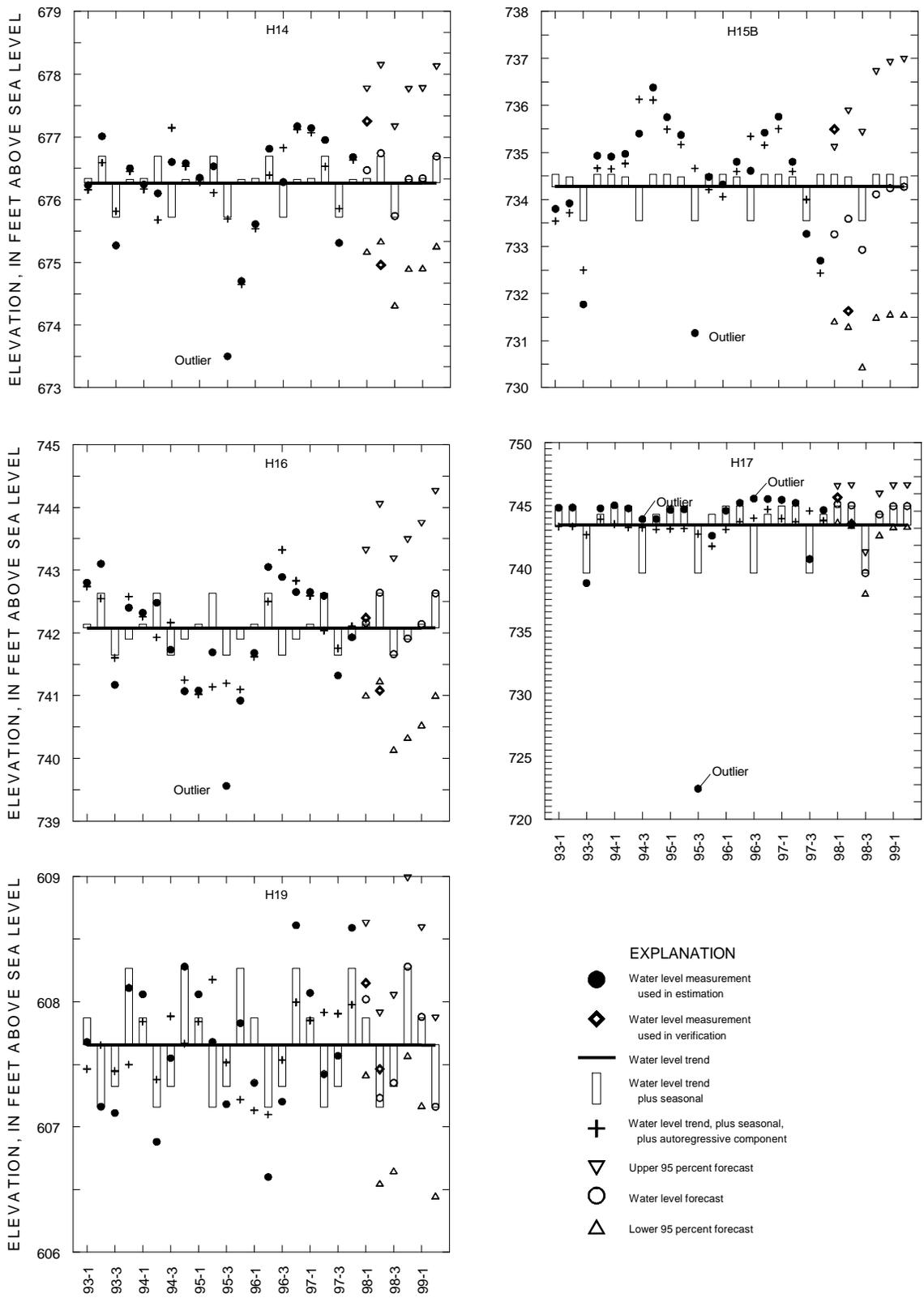


Figure 5. Water levels and time-series components at wells H14, H15B, H16, H17, and H19.

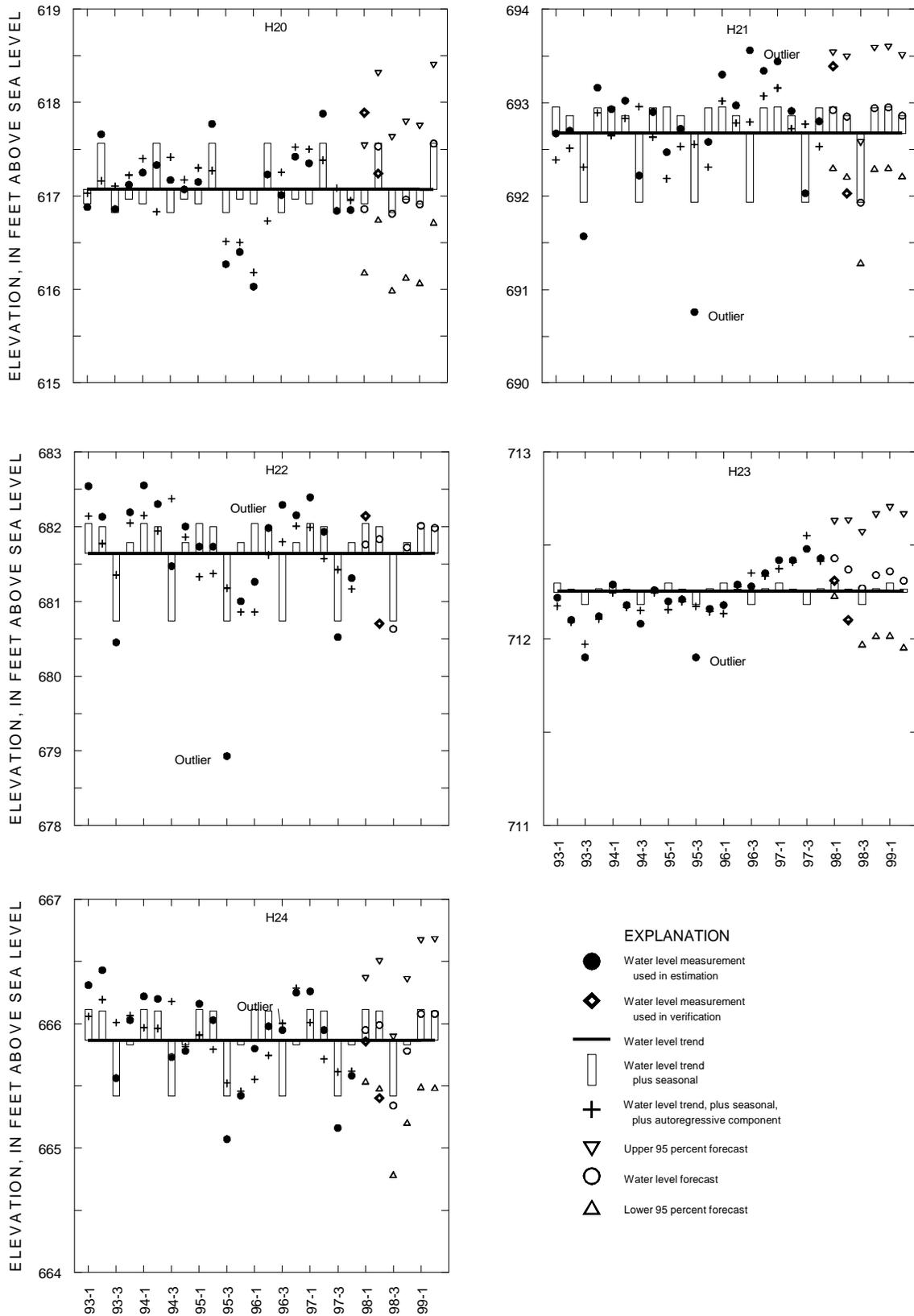
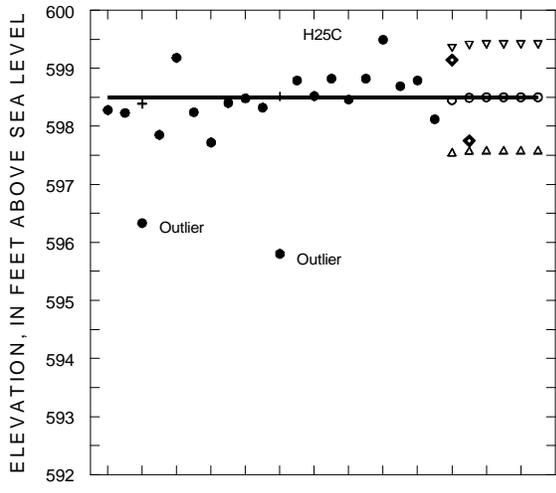
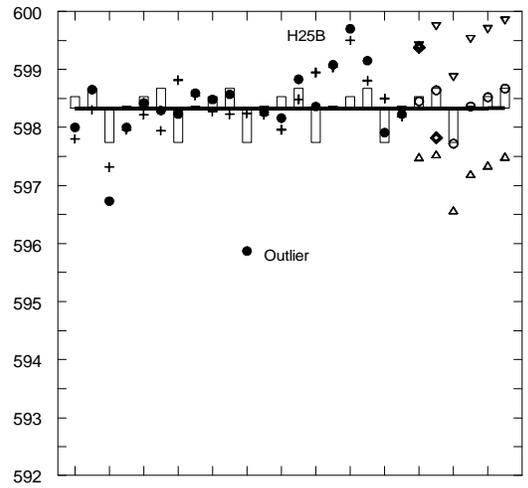
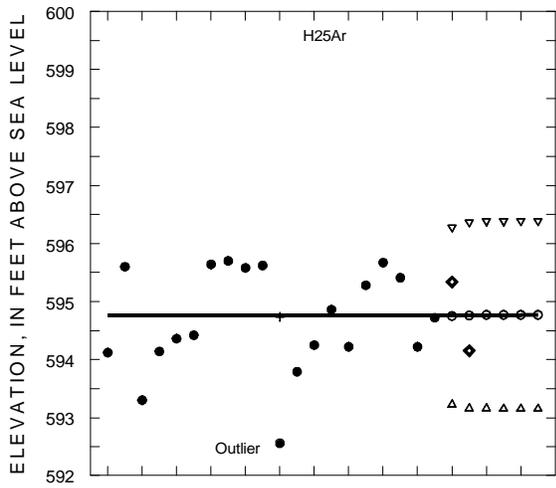


Figure 6. Water levels and time-series components at wells H20, H21, H22, H23, and H24.



- EXPLANATION**
- Water level measurement used in estimation
 - ◇ Water level measurement used in verification
 - Water level trend
 - Water level trend plus seasonal
 - + Water level trend, plus seasonal, plus autoregressive component
 - ▽ Upper 95 percent forecast
 - Water level forecast
 - △ Lower 95 percent forecast

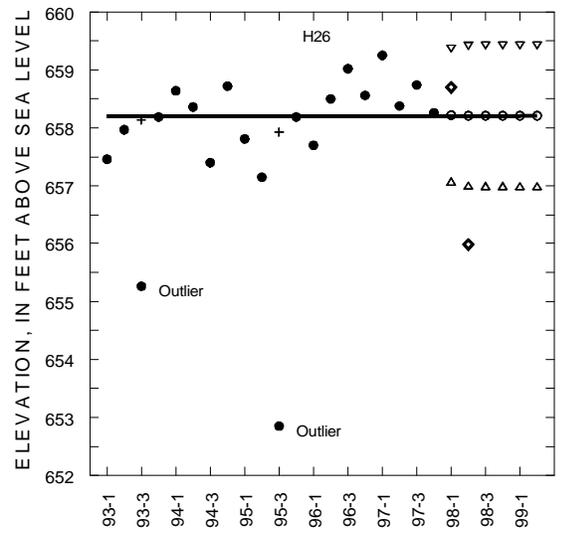
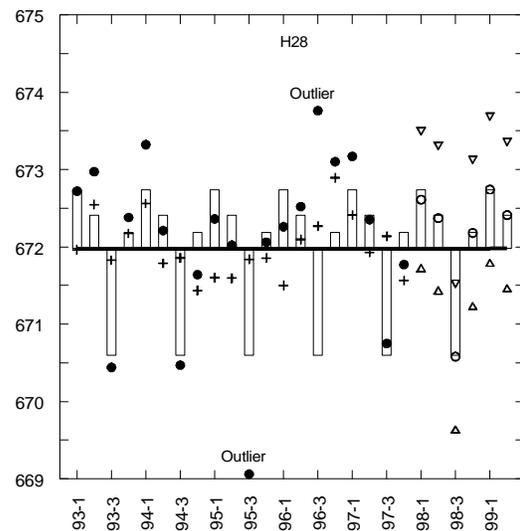
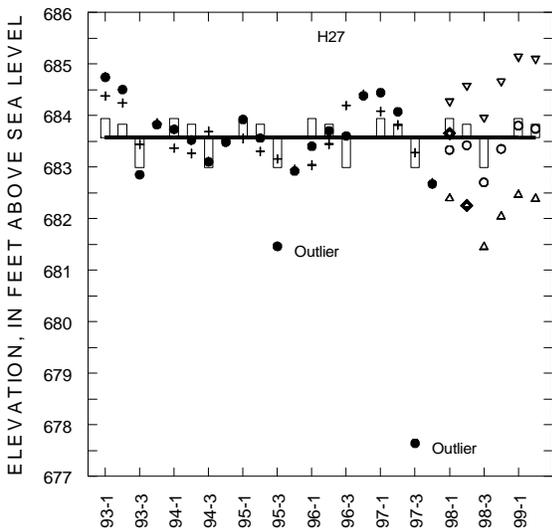


Figure 7. Water levels and time-series components at wells H25Ar, H25B, H25C, and H26.



- EXPLANATION**
- Water level measurement used in estimation
 - ◆ Water level measurement used in verification
 - Water level trend
 - Water level trend plus seasonal
 - + Water level trend, plus seasonal, plus autoregressive component
 - ▽ Upper 95 percent forecast
 - Water level forecast
 - △ Lower 95 percent forecast

Figure 8. Water levels and time-series components at wells H27 and H28.

wells), indicating close agreement with the forecasts. In contrast, during the second quarter of 1998, only 65.4 percent of the measured water-levels were within the 95 percent of the lead-2 forecast intervals. All measured water levels that were outside the forecast intervals were too low. Part of the discrepancy between measured and forecast intervals during the second quarter may be associated with low second-quarter precipitation amounts. Specifically, during the second quarter of 1998, total precipitation was 5.62 inches, compared with an average second quarter precipitation of 9.14 inches during the model estimation (1993-1997) period. Identifying a measure of precipitation that is more highly related to the water levels than quarterly totals may help improve the forecast model performance.

SPATIAL INTERRELATIONS AMONG GROUND-WATER LEVEL MEASUREMENTS

In addition to the temporal components of ground-water measurements, ground-water levels also are spatially correlated. This correlation is thought to be positive, that is, where point measurements indicate above (below) average water levels, nearby points are also likely to be above (below) average. The physical basis for this assumption results from the regional nature of rainfall, which is a primary cause of ground-water level fluctuations (in areas not influenced by pumping), and from the ground-water flow response to water-level gradients.

The positive correlation among water-level measurements creates the potential for data redundancy and inefficiency in monitoring networks. Quantifying the spatial correlation structure is needed to assess its impact on monitoring effectiveness. Variogram analysis (Cressie, 1991, and Isaaks and Srivastava, 1989) is a common technique for estimating the spatial correlation structure as a function of separation distances between measurements. It is, however, not particularly appropriate for analysis of monitoring networks where there are few unique separation distances because repeated measurements are made at a limited number of wells, or where separation distance itself is not a reliable indicator of spatial correlation because of differences in the hydraulic properties of geologic units separating wells in different

strata. Thus, the analysis of spatial interrelations developed in this report does not rely on a consistent relation between separation distance and correlation to analyze effectiveness, but utilizes the available statistical evidence directly. It is similar to variogram analysis, however, in that an empirical covariance structure is used as a basis for a more theoretically appropriate model covariance structure.

In this report, the covariance matrix is used to describe how water levels vary together. For this square matrix, the number of rows (columns) corresponds to the number of wells in the network. Variances of water-level components at individual wells are contained on the main diagonal of the matrix, with covariances between water-level components at different wells on off-diagonal terms. Simple correlation coefficients can be computed by dividing individual covariances by the square root of the product of corresponding variance components. The covariance matrix is symmetric along the main diagonal.

To interpret the covariance matrix for network effectiveness, each water-level sequence was adjusted to remove any nonstationarity so that water-level measurements represented the same population. Thus, trend and seasonal components, when present, were subtracted from the values of the water-level measurements, resulting in residual sequences that had a mean of zero. The effects of outliers also were removed so that the data analyzed were approximately normally distributed. About 93.3 percent of the measurements are thought to arise from a normal probability distribution, for which the model is applicable; 6.7 percent of the measurements, corresponding to outliers, may arise from a different distribution, which is not accounted for by the model. Finally, effects of temporal correlation in the residual sequences were removed by filtering. Filtering, as used in this report, is a mathematical operation that subtracts the product of the autoregressive coefficient and the value of the lag 1 residual (at the previous time step) from the residual value, resulting in a loss of one measurement value in the analysis. The resulting filtered residual sequences are referred to as innovation sequences ϵ_t (eq. 4) in this report.

An empirical estimate of the covariance of the innovation sequences was computed as a prelimi-

nary estimate of the structure of the spatial interrelations. Specifically, row vectors of innovations of length $n=19$ (5 years of quarter measurements minus the measurement used in filtering) were stacked for each of the $p=26$ monitoring wells to form the $\tilde{\epsilon}_{pxn}$ matrix. Then, the empirical covariance matrix $S_{\tilde{\epsilon}}$ was computed as:

$$S_{\tilde{\epsilon}} = \frac{1}{n-1} \tilde{\epsilon}_{pxn} \tilde{\epsilon}'_{pxn}. \quad (5)$$

Results of this calculation indicate that the signs of elements in the empirical covariance matrix were generally positive, although a few were negative (fig. 9). Because of the positive spatial correlation in rainfall patterns and the flow response of ground-water levels to head gradients, however, the negative covariances are not physically plausible. More likely, the negative covariance estimates are the result of sampling error (resulting from a limited number of measurements rather than from errors in the measured values). Similarly, the statistical significance of positive covariances are difficult to evaluate. To overcome these limitations for inferring spatial interrelations from an empirical covariance matrix, a model was developed for the covariance structure.

A Graphical Model of the Covariance Structure

Graphical modeling is a form of multivariate analysis that represents statistical interrelations by use of independence graphs (Edwards, 1995, Whittaker, 1990). Graphical models are based on an analysis of the inverse empirical covariance matrix, $\hat{\Omega}_Y = S_Y^{-1}$. Off-diagonal elements of $\hat{\Omega}_Y$ (referred to as the precision matrix) are the estimated partial correlation coefficients. Under a multivariate normal assumption, any two variables (water-level innovations) are conditionally independent if the corresponding elements in the inverse covariance matrix are equal to zero. Thus, replacing elements in the precision matrix with zero that are not significantly different from zero, simplifies the description of spatial interrelations.

In this analysis, graphical models were developed by initially assuming an independence model¹, that is, all partial correlations were zero and the corresponding initial model covariance matrix, $\tilde{\Sigma}_0$, was equal to the diagonal matrix of $S_{\tilde{\epsilon}}$. Then, a set of 26 alternative covariance matrices

$\tilde{\Sigma}_{\{1\}}$ was created corresponding to the covariance structures for all possible single-well monitoring networks. Elements in the alternative covariance matrices were computed in an iterative manner to meet the constraints that (1) the diagonal elements in the covariance matrix set $\tilde{\Sigma}_{\{1\}}$ were equal to the diagonal elements of $S_{\tilde{\epsilon}}$, (2) the off-diagonal element pairs in $\tilde{\Sigma}_{\{1\}}$ were equal to the corresponding elements in $S_{\tilde{\epsilon}}$, and (3) the off-diagonal element pairs in the precision matrix set $\tilde{\Omega}_{\{1\}}$ were the only non-zero partial correlation pairs in the precision matrices. The set of maximized log-likelihood functions was evaluated for the set of covariance and precision matrices as:

$$\hat{L}_{\{1\}} = \frac{-nq_{\{1\}} \ln(2\pi)}{2} + \frac{-n \ln(|\tilde{\Sigma}_{\{1\}}|)}{2} - \frac{nq_{\{1\}}}{2}, \quad (6)$$

where

n equals the number of innovations,

$q_{\{1\}}$ is the trace (sum of the diagonal components) of the matrix product $\tilde{\Omega}_{\{1\}} \tilde{\Sigma}_{\{1\}}$, and

$\tilde{\Sigma}_{\{1\}}$ is the determinant of the alternative covariance matrix.

A set of Bayesian Information Criteria (*BIC*) values was computed as

$$BIC_{\{1\}} = -2 \ln L_{\{1\}} - \ln(n)p_{\{1\}} \quad (7)$$

where $p_{\{1\}}$ is the number of model parameters.

The maximum value of the BIC_j in the set $BIC_{\{1\}}$ and the corresponding monitoring well was selected as the most effective single-well monitoring network. The graphical modeling analysis proceeded in a stepwise manner to identify additional effective wells until $BIC_{\{k\}}$ could not be decreased any further.

An independence graph resulting from this analysis is interpreted as follows: (1) vertices (corresponding to water-level innovations at individual wells) that are not connected by an edge are not significantly correlated (assuming multivariate normality, they are independent); (2) vertices that are directly connected by an edge are correlated (not independent), and (3) vertices that are con-

1. A saturated model, in which all partial correlations were initially assumed to be significant, could not be estimated because of the limited number of time series measurements.

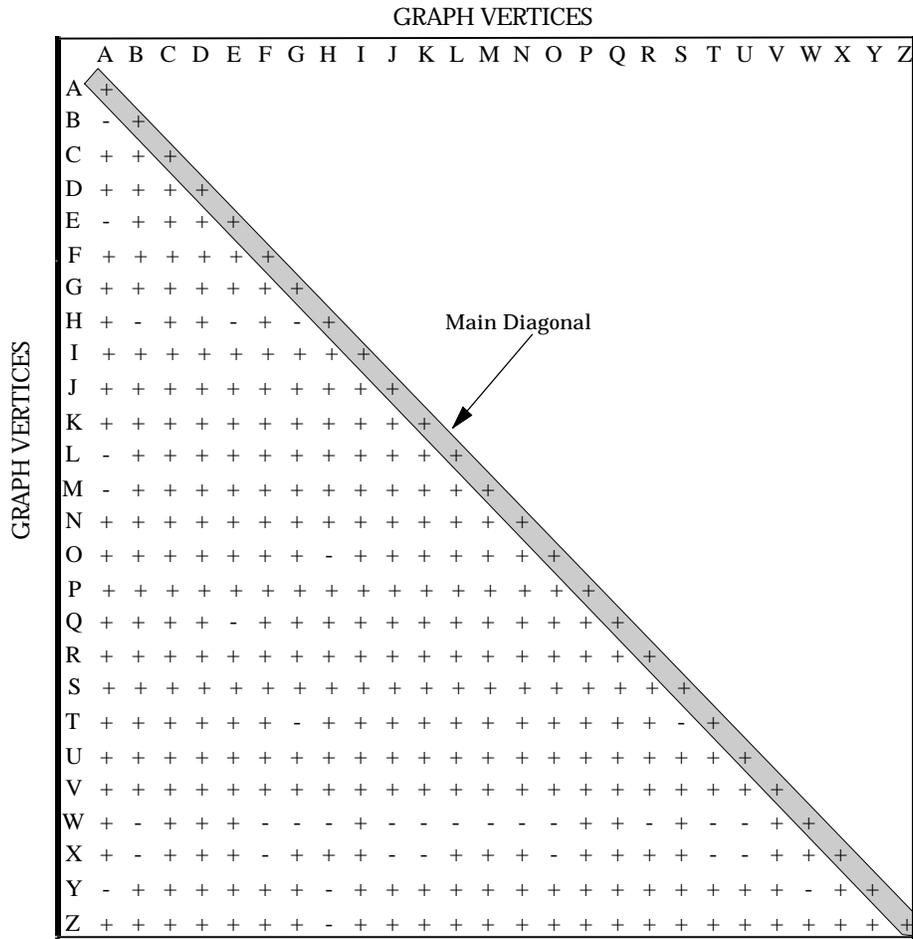


Figure 9. Signs of elements of the empirical covariance matrix (Note: the covariance matrix is symmetrical about the main diagonal, so only the lower triangular portion is shown).

nected through an intervening vertex are uncorrelated after accounting for the effect of the intervening vertex (conditionally independent). Thus, a diagram of the independence graphs depicts the significant spatial interrelations among water-level innovations and is associated with a model covariance matrix.

Results of Graphical Modeling

A graphical model of water-level innovations was developed by use of MIM 3.0 (HyperGraph Software, 1998). Starting with an independence model, a forward selection process of 29 steps reduced the Bayesian Information Criteria from 600.3 for BIC_0 to 292.0 for BIC_{29} (fig. 10). The selected model included 23 sets of mutually connected sets of vertices (cliques) formed by 29 of the possible 325 edges (fig 11). Of the 23 cliques, 19 consisted of two-vertex sets, such as $\{A I\}$ and $\{B K\}$, and 4 were three-vertex sets, including $\{C J K\}$ and $\{E F K\}$. Of the 26 vertices, 11 were first degree (directly connected to only 1 other vertex), 5 were second degree, 6 were third degree, 1 was fourth degree, and 3 were fifth degree vertices. The degree of a vertex indicates the number of conditioning vertices needed for independence at that vertex. Thus 11 wells are independent of innovations throughout the remaining network, given water levels at only one other well. The water-level monitoring network is connected, in the sense that there is a path (a sequence of edges) between every pair of vertices.

Three principal sub-networks can be defined on the basis of separation by vertex L (Well H15B): N_A , defined as the set of vertices, equal to $\{A B C F H I J K M N Q R T U\}$, N_B equal to $\{D S V W X Z\}$, and $N_C = \{E G O P Y\}$. This separation implies that water-level innovations in any one sub-network are conditionally independent of innovations in the remaining two subnetworks given data at well H15B. The spatial interrelations indicated by the graphical model are shown in fig. 12, where the vertices are located geographically within Huron County.

ACHIEVING EFFECTIVENESS BY RANKING MONITORING WELLS AND SPECIFYING NETWORK CONSTRAINTS

Time-series analyses were used to identify regular and irregular components of water-level fluctua-

tations in monitoring wells. The regular components, such as trend and seasonal characteristics, were used to describe and forecast water levels. The irregular component, referred to as the autoregressive component, was used to characterize the uncertainty in water-level information. After filtering to remove the influence of autocorrelation, the irregular component was used with graphical modeling to quantify the spatial interrelations among water-levels. This description of spatial interrelations provides a basis for improving the monitoring effectiveness. Specifically, the description can be used to identify an appropriate subset of monitoring wells to meet specifications on network uncertainty or constraints on monitoring resources.

In this report, the uncertainty of the network refers to the quarterly variability in water-level innovations that are not accounted for by monitoring. If quarterly monitoring were to continue at all wells, the uncertainty of the network would be zero. If monitoring were discontinued, an estimate of uncertainty could be based on the sum of diagonal elements (trace) of the model covariance matrix. For all other monitoring options, an estimate of the network uncertainty could be based on those variance components that are not accounted for by monitored wells. The square root of the summed variance components is used here to provide a measure of uncertainty that is consistent with the measurement unit of feet.

The model covariance matrix can be partitioned into four components to correspond with a set of unmonitored wells $\{Y\}$ and a set of monitored wells $\{Z\}$ as

$$\tilde{\Sigma}_{29} = \begin{bmatrix} \tilde{\Sigma}_{\{YY\}} & \tilde{\Sigma}_{\{ZY\}} \\ \tilde{\Sigma}_{\{YZ\}} & \tilde{\Sigma}_{\{ZZ\}} \end{bmatrix}, \quad (8)$$

where

$\tilde{\Sigma}_{29}$ is the model covariance structure inferred from graphical modeling,

$\tilde{\Sigma}_{\{YY\}}$ is the model covariance among the unmonitored wells,

$\tilde{\Sigma}_{\{ZZ\}}$ is the model covariance among the monitored wells, and

$\tilde{\Sigma}_{\{YZ\}}$, and its transpose $\tilde{\Sigma}_{\{ZY\}}$, is the model covariance among the monitored and unmonitored wells.

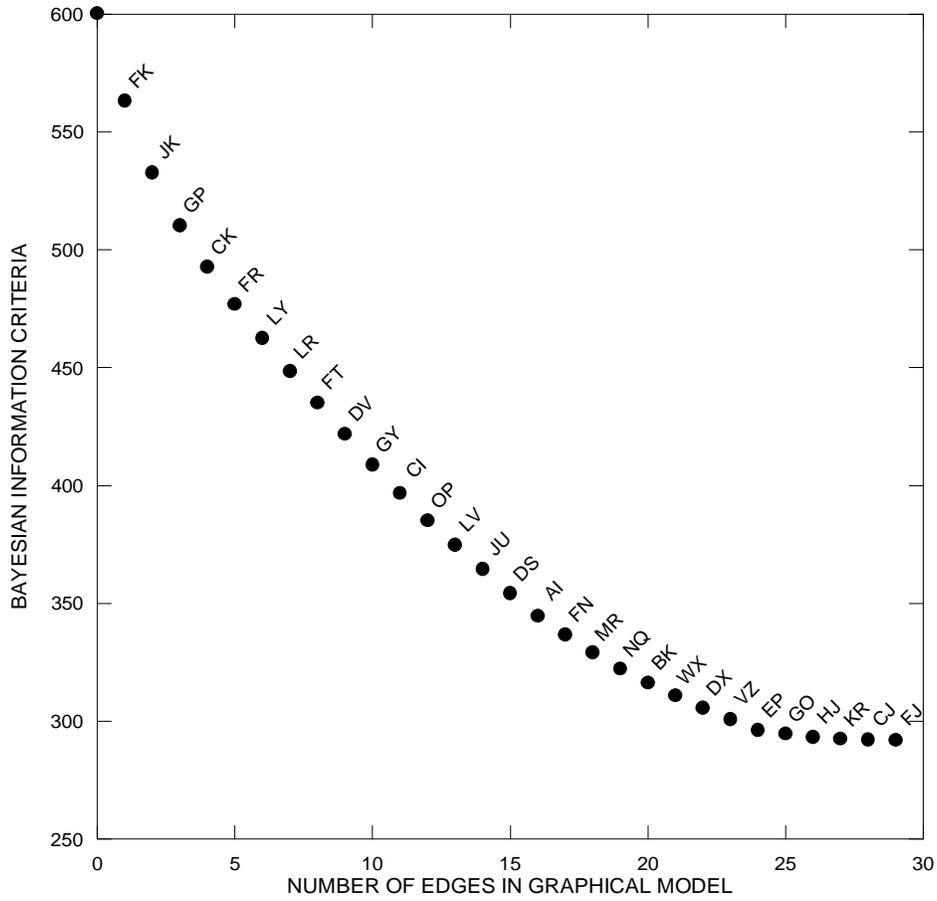


Figure 10. Relation between the number of edges in the graphical model and the Bayesian Information Criteria.

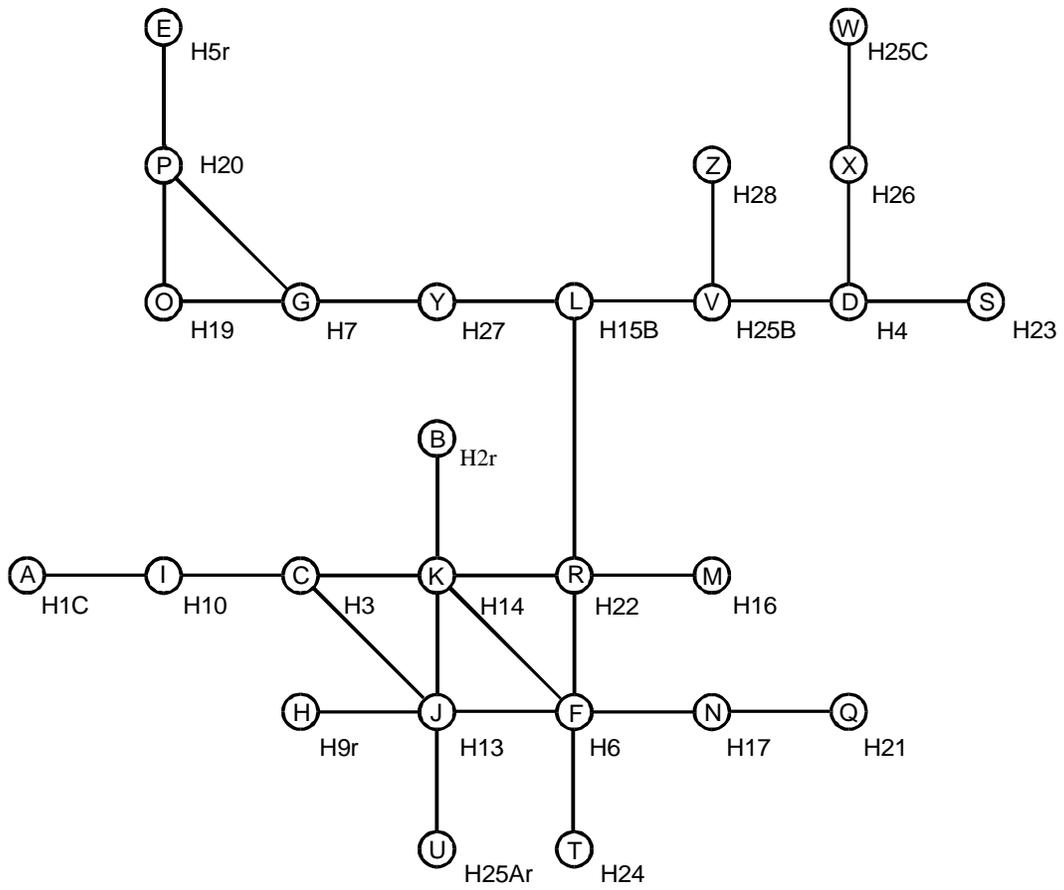


Figure 11. Independence graph of water-level innovations at monitoring wells.

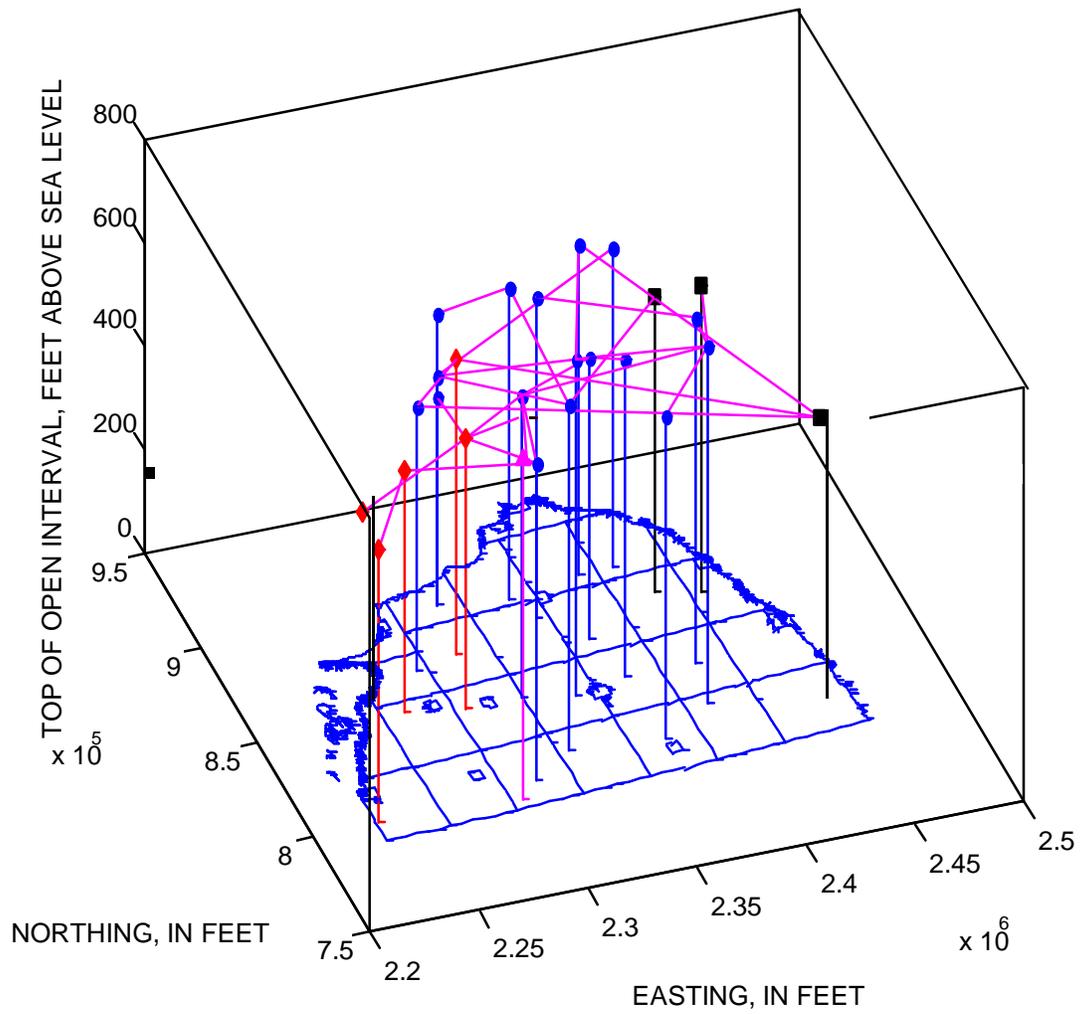


Figure 12. Geographic reference for the independence graph in Huron County, Michigan.

EXPLANATION

Vertices

- Coldwater unit
- Marshall aquifer
- ▲ Glaciofluvial aquifer
- ◆ Saginaw aquifer

Edge

- Indicates conditional dependence

This partitioning strategy is used to describe the network uncertainty remaining after selecting a subset of monitoring wells $\{Z\}$ of varying size by computing the model partial covariance structure as:

$$\tilde{\Sigma}_{YY \cdot Z} = \tilde{\Sigma}_{YY} - \tilde{\Sigma}_{YZ} \tilde{\Sigma}_{ZZ}^{-1} \tilde{\Sigma}_{ZY}, \quad (9)$$

where

$\tilde{\Sigma}_{YY \cdot Z}$ is the model partial covariance structure of the set of unmonitored wells $\{Y\}$, given that wells in the set $\{Z\}$ are being monitored, and $\tilde{\Sigma}_{ZZ}^{-1}$ is the inverse model covariance of the monitored wells.

The maximum uncertainty of water-level innovations in the network corresponds to having no monitoring wells, that is

$tr(\tilde{\Sigma}_{29}) = tr(\tilde{\Sigma}_{YY \cdot Z})$ for $\{Z\} = \{\emptyset\}$, where tr is the trace. From this point, a set of 26 partial covariance matrices $\tilde{\Sigma}_{YY \cdot Z_{\{1\}}}$ were computed corresponding to all 26 single-well networks. The most effective single-well monitoring network was identified by selecting the partial covariance matrix with the minimum trace of $\tilde{\Sigma}_{YY \cdot Z_1}$. The well corresponding to this network was added to the set $\{Z\} = \{Z_1\}$. The computation, selection, and addition procedure was repeated for the 25 two-well networks that are possible after the first well is selected. The computation proceeded in a similar manner until all 26 wells were included in the set of monitored wells.

The sequence of wells identified provides a ranking of monitoring wells from the most effective (well H6) to the least effective (well H7) (fig. 13). The nonlinear decrease in uncertainty shows that water-level innovations are not equally (co)variable at all wells. Further, the ranking provides a mechanism for selecting the most effective subset of wells for meeting uncertainty specifications or monitoring constraints.

For example, if monitoring resources were sufficient to operate only 6 monitoring wells, the most effective network would be the set $\{H6, H5, H15B, H25Ar, H26, H9r\}$ (fig. 13). This is the same set of wells that most effectively accounts for 50 percent of the uncertainty currently monitored by the network of 26 wells. By extension, figure 13 can be used to identify the subset of wells that most effectively meets a specified uncertainty for any reduced sized network.

The relationship in figure 13 is applicable only to the proportion (93.3 percent) of the measurements likely to arise from the multivariate normal distribution described by the graphical model. That is, about 6.7 percent of water-level measurements were classified as outliers that arise from an unknown probability distribution. The network uncertainty from these outlier-type measurements are not explicitly accounted for in the analysis. If the probability of outliers were equal at all wells, however, the most effective network would consist of the same subset of wells, although the percentage of total uncertainty accounted for by a reduced network would be less than that indicated by figure 13.

SUMMARY

Quarterly ground-water level measurements between 1993 and 1998 at 26 wells in Huron County, Michigan were analyzed to assess the effectiveness of the monitoring network. A structural time-series analysis identified trend, seasonal, and autoregressive components and statistical outliers in the series at individual wells. Results indicate that trend components at all sites were constant levels and that residuals contained an autoregressive component. In addition, fluctuations of water levels in 22 wells followed a fixed seasonal pattern. A total of 35 outliers were identified, 21 of which occurred in third quarter of 1995. Although the large number of outliers in 1995 may be related to unusual rainfall during that year, the number of outliers may indicate that sampled water levels arise from two statistical distributions: a normally distributed population constituting perhaps 93.3 percent of the measurements and a more extreme distribution generating the remaining 6.7 percent.

The structural time-series models were used to forecast water-level intervals for 6 quarters beginning with the first quarter of 1998. Comparison between forecast intervals and measurements during the first two quarters of 1998 produced somewhat discrepant results. Results for the first quarter indicate that forecast intervals were in close agreement with measured values; results in the second quarter indicate that forecast intervals were often higher than measured values. Poor results during the second quarter of 1998 may be related to unusually low precipitation amounts. Finally, the

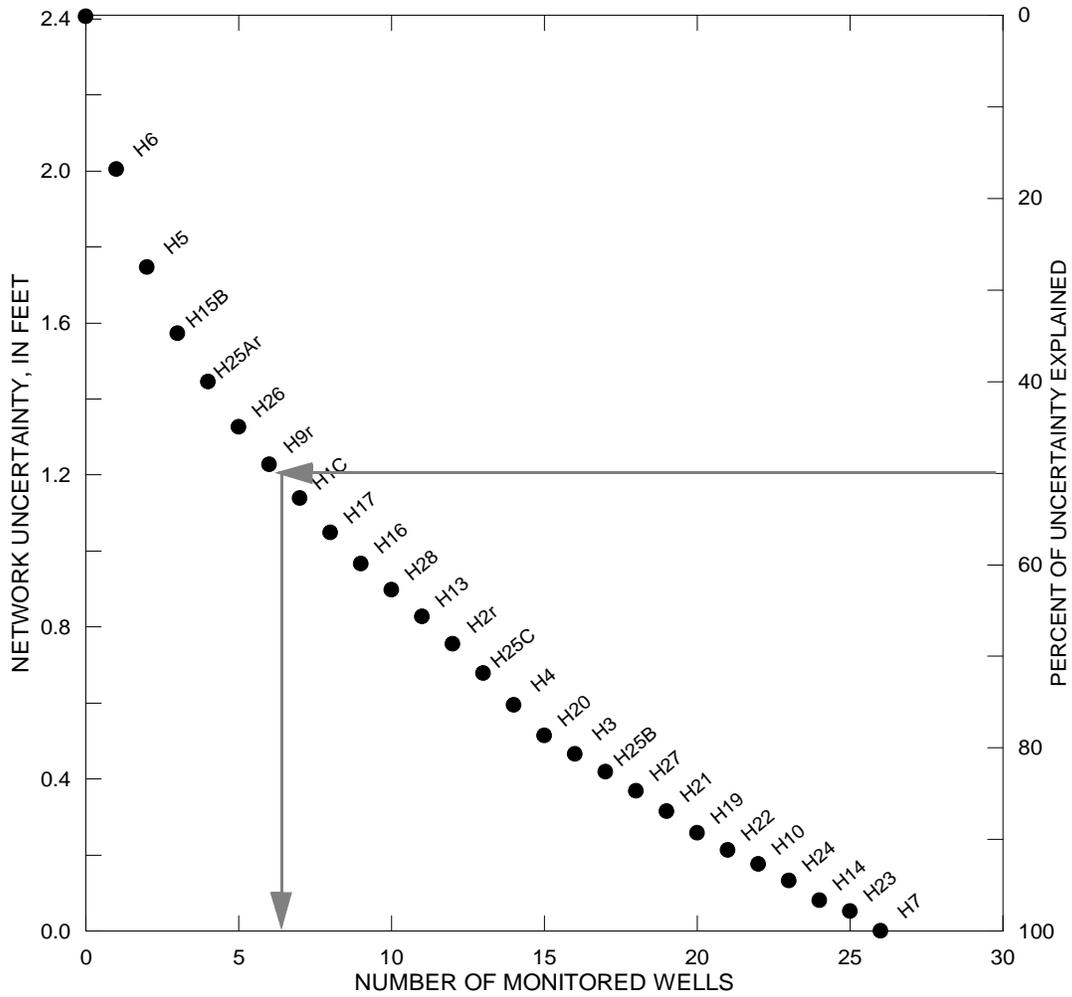


Figure 13. Relation between number of wells monitored and network uncertainty.

autoregressive coefficient estimated in the structural time series models were used to filter the autoregressive errors and produce innovation sequences used in the evaluation of the monitoring effectiveness.

The innovation sequences were used to compute an empirical estimate of the spatial interrelations among water-level innovations. Although the spatial interrelations were expected to be positive, some of the individual empirical innovation covariances were negative. Although the negative empirical covariances are thought to arise from sampling error, further analysis was needed to determine the statistically significant interrelations. Thus graphical modeling, which is a form of multivariate analysis, was used to model the spatial interrelations among water level innovations. The graphical model identified 29 significant interrelations among the possible 325 estimated by the empirical covariance matrix. Thus the graphical model not only simplified the description of interrelations among wells, but also resulted in a model covariance matrix in which no covariance elements were less than zero.

The model covariance matrix was used to rank the effectiveness of alternative subsets of monitoring wells. The uncertainty of a network with no monitoring wells was computed as the square root of the trace of the model covariance matrix. For sub-networks composed of from 1 to 25 monitoring wells, effectiveness was based on a partitioning of the model covariance matrix. The uncertainty for the full network was assumed to be zero. The results provide a mechanism for selecting subsets of monitoring wells, should reductions in the monitoring network be required.

REFERENCES CITED

- Box, G.E.P., and Jenkins, G.M., 1976, Time series analysis forecasting and control: Oakland, California, Holden-Day, 575 p.
- Brockwell, P.J., and Davis, R.A., 1987, Time series: theory and methods: New York, Springer-Verlag, 519 p.
- Cressie, N.A., 1991, Statistics for spatial data: New York, A Wiley-Interscience Publication, John Wiley, 900 p.
- Edwards, David, 1995, Introduction to graphical modeling: New York, Springer-Verlag, 274 p.
- Harvey, A.C., 1994, Forecasting, structural time series models and the Kalman filter: Cambridge University Press, 554 p.
- HyperGraph Software, Inc., 1998, MIM 3.0: Bymarken 38, DK-4000, Roskilde, Denmark
- Isaaks, E.H., and Srivastava, R.M., 1989, Applied geostatistics: New York, Oxford University Press, 561 p.
- Koopman, S.J., Harvey, A.C., Doornik, J.A., and Shephard, Neil, Stamp 5.0--Structural time series analyser, modeller and predictor: London, Chapman and Hall, 382 p.
- Sweat, M.J., 1991, Hydrogeology of Huron County, Michigan: U.S. Geological Survey Water-Resources Investigations Report 91-4133, 68 p.
- , 1995, Ground-water levels in Huron County, Michigan, March 1993 through December 1994: U.S. Geological Survey Open-File Report 95-430, p. 9 p.
- , 1996, Ground-water levels in Huron County, Michigan, January 1995 through December 1995: U.S. Geological Survey Open-File Report 96-175, p. 8 p.
- , 1997, Ground-water levels in Huron County, Michigan, January 1996 through December 1996: U.S. Geological Survey Open-File Report 97-196, p. 9 p.
- Whittaker, Joe, 1990, Graphical models in applied multivariate statistics: New York, John Wiley and Sons, 448 p.