AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Water Supply

1976

UNIVERSITY OF CALIFORNIA
DAVIS

SEP 27 1976

Agricultural Economics Library

ALTERNATIVE MODELS FOR ESTIMATING THE TIME SERIES COMPONENTS OF WATER CONSUMPTION DATA

by

Hiroshi Yamauchi and Wen-yuan Huang*

For presentation at the American Agricultural Economics Association meeting, at Pennsylvania State University, 15-18 August 1976.

*Associate Professor, Department of Agricultural and Resource Economics and the Water Resources Research Center, University of Hawaii; Assistant Professor, Department of Agricultural and Resource Economics, University of Hawaii, respectively.

ABSTRACT

Additive and multiplicative models are used to analyze the trend, cyclical, seasonal, and irregular components of water consumption time series data. Evaluation of these components opens up important opportunities for improving on existing water management practices and for better evaluating alternative policy measures.

NEED FOR BETTER UNDERSTANDING OF WATER CONSUMPTION PATTERNS

Better understanding of the empirical nature of water demand has become an important prerequisite for further improvements in planning and management of water resources. The significance of this statement has recently come into clear focus with the findings of Lobb (1975) with regard to the high costs of extracapacity typically found in water supply systems to meet peak load demands.

Recently completed studies of water consumption patterns and trends in the Honolulu Board of Water Supply (BWS) service areas serve as a useful starting point to apply modern statistical techniques for improving the analysis of water consumption data (Oh and Yamauchi 1974; Moncur 1975). This in turn can be very helpful in making short- and long-run projections for water supply requirements, and also in generating timely information for such policy-related matters as pricing, repayment, cost distribution, and the like.

The underlying statistical principles involved are applicable to water consumption data and involve procedures for identification and postulation of statistical models, estimation of parameters, and diagnostic checks as to the adequacy of the results.

General Model

Typically, the straightforward display of aggregate time series data shows a rising pattern for seasonal fluctuations which vary over time (Fig. 1). In general, the behavior of such time series data reflects the combined effects of many complex factors which influence the trend, cyclical, seasonal, and irregular components of the series. Even if the underlying factors are not fully understood, these separate components can be hypothesized to exist after careful examination of the data. Thus, without any further specification of the nature of the relationships that may be involved, a general form of the model can be represented as follows:

$$Q_t^m = f(T, C, S, I)$$

Where:

Q_t^m = average daily water consumption in month m and time interval t

T = trend component

C = cyclical component

S = cyclical component

I = irregularities

Trend (T) represents all those factors that influence the general long run tendency of the series. If there are factors that cause the series to move in relatively long and uneven swings, much like the gradual expansions and contractions in the business cycle and/or long-run climatic patterns, such factors can be captured by the cyclical (C) component. For convenience in long-run forecasting, both the trend and cyclical components (TC) can be measured jointly with a low-order polynominal in t, the time interval.

The seasonal (S) component represents those factors which influence the series to fluctuate in a more or less regular patterns within the year. Such intra-year fluctuations are primarily the result of major shifts in the weather pattern from summer to winter. This seasonal component for water consumption data can be efficiently measured by a dummy variable technique which assigns either a 1 or 0 to each time interval depending upon whether or not that interval falls on a particular month of the year, i.e., January to December.

The irregular (I) component is essentially the residual or whatever variations in the time series data that is not explained systematically by the previous T, C, and S components. Such irregularities may occur as a result of a wide variety of factors, such as unusual weather conditions, unexpected water leakages, measurement and/or specification errors, etc.

This general formulation can be further specified in either additive or multiplicative terms depending upon the nature of the relationships between the various components. If the many underlying factors are independent in their influences on the separate components, then an additive function is appropriate, as

$$Q_t^m = T_t + C_t + S_t + I_t$$

On the other hand, if these multiple influences are interdependent with each other, then a multiplicative function is more appropriate:

$$Q_t^m = T_t \cdot C_t \cdot S_t \cdot I_t$$

In either case, in order to statistically analyze the observed data, these models require further specifications in operational terms.

Explicit Models

Several possibilities exist for the identification of explicit mathematical forms. The simplest is of course the straight additive model which does not require any data transformations in its application. More complex multiplicative models can also be specified in either nonexponential or exponential forms. The nonexponential form, however, presents an operational problem. Dummy variables (1, 0) are useful for measuring monthly (or seasonal effects) and their log transformations would result in 0 and $-\infty$ values, respectively. For practical purposes, therefore, only the straight additive (eq. 1) and the exponentially multiplicative (eq. 2) models are explored.

1. Additive Model:

$$Q_t^m = \alpha_0 + \sum_{i=1}^{4} \alpha_i \, t^i + \sum_{m=1}^{11} \beta_m X_m + U_t^m$$

2. Exponentially Multiplicative Model:

$$Q_t^m = \exp(\alpha_0 + \sum_{i=1}^{4} \alpha_i \, t^i + \sum_{m=1}^{11} \beta_m X_m + U_t^m)$$

Where:

Q_t^m = average daily water consumption (in 10,000 gals/day)

m = month of year
0 = January (base month)
1 = February
2 = March
3 = April
.
.
.
11 = December

t = time interval (month) with December 1960 as base
1-12 = January-December 1961
13-24 = January-December 1962
.
.
.
180-187 = January-July 1975

$\alpha_0 + \sum_{i=1}^{4} \alpha_i \, t^i$ = polynomial representing combined trend and cyclical components, i.e., TC

α = regression coefficients
t = time intervals (months)

$\sum_{m=1}^{11} \beta_m X_m$ = expression of dummy variables representing seasonal (S) components

β = regression coefficients
X = 1,0 dummy variables for month of year with January as base month

U_t^m = error term representing irregular (I) component, assumed to be $N(0, \sigma^2)$

The dummy variables expression for seasonality effects makes it possible to simultaneously estimate and study the water consumption behavioral patterns for the different months of the year over the 14 years of record.

That is, for

$$\text{January:} \quad Q_t^0 = \alpha_0 + \sum_{i=1}^{4} \alpha_i t^i + e_t^0$$

$$\text{February:} \quad Q_t^1 = \alpha_0 + \sum_{i=1}^{4} \alpha_i t^i + \beta_1 X_1 + e_t^1$$

$$\vdots \qquad\qquad \vdots$$

$$\text{December:} \quad Q_t^{11} = \alpha_0 + \sum_{i=1}^{4} \alpha_i t^i + \beta_{11} X_{11} + e_t^{11}$$

In matrix form, these two operational models appear as follows with the only difference being the transformed log consumption values in the multiplicative model.

MATRIX FORMS:

$$Q = XB + U \text{ (additive)}$$
$$\text{Ln } Q = XB + U \text{ (exponentially multiplicative)}$$

$$
\begin{pmatrix}
Q_1^0 \\ Q_2^1 \\ \vdots \\ Q_{12}^{11} \\ \hline Q_{13}^0 \\ Q_{14}^1 \\ \vdots \\ Q_{24}^{11} \\ \hline \vdots \\ \hline Q_{169}^0 \\ Q_{170}^1 \\ \vdots \\ Q_{180}^{11} \\ \hline Q_{181}^0 \\ Q_{182}^1 \\ \vdots \\ Q_{187}^6 \\ \vdots
\end{pmatrix}
=
\left(\begin{array}{ccccc|ccccccccccc}
1 & 1 & 1^2 & 1^3 & 1^4 & 0&0&0&0&0&0&0&0&0&0&0 \\
1 & 2 & 2 & 2 & 2 & 1&0&&&&&&&&&0 \\
 & & & & & 0&1&&&&&&&&& \\
\vdots & & & & & \vdots&&&&&1&&&&& \\
1 & 12 & 12^2 & 12^3 & 12^4 & 0&0&&&&&&&&&1 \\ \hline
1 & 13 & 13^2 & 13^3 & 13^4 & 0&0&0&0&0&0&0&0&0&0&0 \\
1 & 14 & 14^2 & 14^3 & 14^4 & 1&0&&&&&&&&&0 \\
 & & & & & 0&1&&&&&&&&& \\
\vdots & & & & & \vdots&&&&&1&&&&& \\
1 & 24 & 24^2 & 24^3 & 24^4 & 0&0&&&&&&&&&1 \\ \hline
\vdots & & & & & \vdots&&&&&&&&&& \\ \hline
1 & 169 & 169^2 & 169^3 & 169^4 & 0&0&0&0&0&0&0&0&0&0&0 \\
1 & 170 & 170^2 & 170^3 & 170^4 & 1&0&&&&&&&&&0 \\
 & & & & & 0&1&&&&&&&&& \\
\vdots & & & & & \vdots&&&&&1&&&&& \\
1 & 180 & 180^2 & 180^3 & 180^4 & 0&0&&&&&&&&&1 \\ \hline
1 & 181 & 181^2 & 181^3 & 181^4 & 0&0&0&0&0&0&0&0&0&0&0 \\
1 & 182 & 182^2 & 182^3 & 182^4 & 1&0&&&&&&&&&0 \\
 & & & & & 0&1&&&&&&&&& \\
\vdots & & & & & \vdots&&&&&&&&&& \\
1 & 187 & 187^2 & 187^3 & 187^4 & 0&0&&&&1&&&&&0 \\
\vdots & & & & \vdots & \vdots&&&&&&&&&&\vdots
\end{array}\right)
\begin{pmatrix}
\alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{11}
\end{pmatrix}
+
\begin{pmatrix}
U_1^0 \\ U_2^1 \\ \vdots \\ U_{12}^{11} \\ \hline U_{13}^0 \\ U_{14}^1 \\ \vdots \\ U_{24}^{11} \\ \hline \vdots \\ \hline U_{169}^0 \\ U_{170}^1 \\ \vdots \\ U_{180}^{11} \\ \hline U_{181}^0 \\ U_{182}^1 \\ \vdots \\ U_{187}^6 \\ \vdots
\end{pmatrix}
$$

DISCUSSION OF RESULTS

The stepwise regression method was applied to the 187 data points, each representing average daily water consumption for the period covering January 1961 to July 1975. In the analysis, the seasonal (*S*) variables were forced into the equations while the combined trend and cyclical (*TC*) variables were allowed to enter into the equation according to F-levels (4.00, 3.90) or tolerance insufficient for further stepping.

The estimated results are summarized in Tables 1 and 2 for the additive and multiplicative models, respectively. High R^2 (0.95 and 0.96) and corresponding *F*-ratios (277 and 307) in both instances indicate very good model identifications. Although the slight improvement (actually less than 1%) in the multiplicative model might suggest that the major effects on the combined *TC* and *S* components are for the most part independent of each other, a closer examination of the results indicate strong evidence of important multiplicative effects that should not be overlooked. In both models, the coefficients are computed from data over the entire series. The overall mean for the series, however, is somewhere close to the middle of the period of 187 intervals. Near this vicinity, both models are fairly close to each other. However, differences begin to show up as the estimates move away in both directions from the overall mean (Fig. 1). In earlier years, the estimated seasonal fluctuations of the multiplicative model tends to be less than those of the additive model. Moving in the other direction, toward the later years, the differences are in opposite relation to each other as compared to the earlier years, i.e., the multiplicative model estimates of seasonal fluctuations tend to be greater than those for the additive model. In essence, the difference in estimates between the two models is explained as follows. In the additive model, the *TC* and *S* components are independent of each other and, therefore, these components are simply added together to arrive at the estimated results. In the multiplicative model, however, the *TC* and *S* components are interrelated functionally and, therefore, the estimated results are arrived at through more complex multiplicative computations involving log transformations.

This is better seen in Figure 2 which plots the observed time series and component parts for both the additive and multiplicative models. Although at first glance there appears to be little or no difference between these plots, in fact there are important interpretive differences. In the additive model, the values are expressed in terms of 10,000 gal/month, whereas in the multipli-

Table 1. Regression Results of Additive Model

Multiple R	0.9768
Multiple R-Square	0.9542
Std. Error of Est.	5.7729

Analysis of Variance

	Sum of Squares	DF	Mean Square	F Ratio
Regression	120191.81	13	9245.523	277.428
Residual	5765.4961	173	33.32655	

Variables in Equation

Variable		Coefficient	Std. Error of Coeff	Std. Reg. Coeff	F to Remove	Level
(y-intercept		48.239)				
B1	3	1.795	2.041	0.019	0.774	1
B2	4	3.487	2.041	0.038	2.918	1
B3	5	7.072	2.041	0.076	12.004	1
B4	6	10.787	2.041	0.116	27.924	1
B5	7	20.085	2.042	0.216	96.793	1
B6	8	25.013	2.042	0.270	150.071	1
B7	9	25.618	2.075	0.268	152.377	1
B8	10	22.774	2.075	0.238	120.441	1
B9	11	12.374	2.075	0.130	35.557	1
B10	12	5.154	2.075	0.054	6.168	1
B11	13	1.158	2.075	0.012	0.311	1
a2	14	0.003	0.000	1.068	602.977	2
a3	17	-0.000	0.000	-0.376	16.418	2

F-Levels (4.000, 3.900) or Tolerance Insufficient for Further Stepping

Table 2. Regression Results and Multiple Model

Multiple R	0.9806
Multiple R-Square	0.9615
Std. Error of Est.	0.0597

Analysis of Variance

	Sum of Squares	DF	Mean Square	F Ratio
Regression	15.318178	14	1.094155	307.140
Residual	0.61273319	176	0.3352402E-02	

Variables in Equation

Variable		Coefficient	Std. Error of Coeff	Std. Reg. Coeff	F to Remove	Level
(y-intercept		3.917)				
a1	1	0.002	0.001	0.348	10.819	2
B1	3	0.019	0.021	0.018	0.757	1
B2	4	0.033	0.021	0.032	2.429	1
B3	5	0.086	0.021	0.083	16.669	1
B4	6	0.122	0.021	0.117	33.166	1
B5	7	0.227	0.021	0.218	115.736	1
B6	8	0.282	0.021	0.270	177.928	1
B7	9	0.288	0.021	0.268	179.718	1
B8	10	0.271	0.021	0.252	159.692	1
B9	11	0.151	0.021	0.140	49.192	1
B10	12	0.069	0.021	0.066	10.210	1
B11	13	0.017	0.021	0.016	0.609	1
a2	14	0.000	0.000	0.752	31.812	2
a3	17	-0.000	0.000	-0.210	8.939	2

F-Levels (4.000, 3.900) or Tolerance Insufficient for Further Stepping

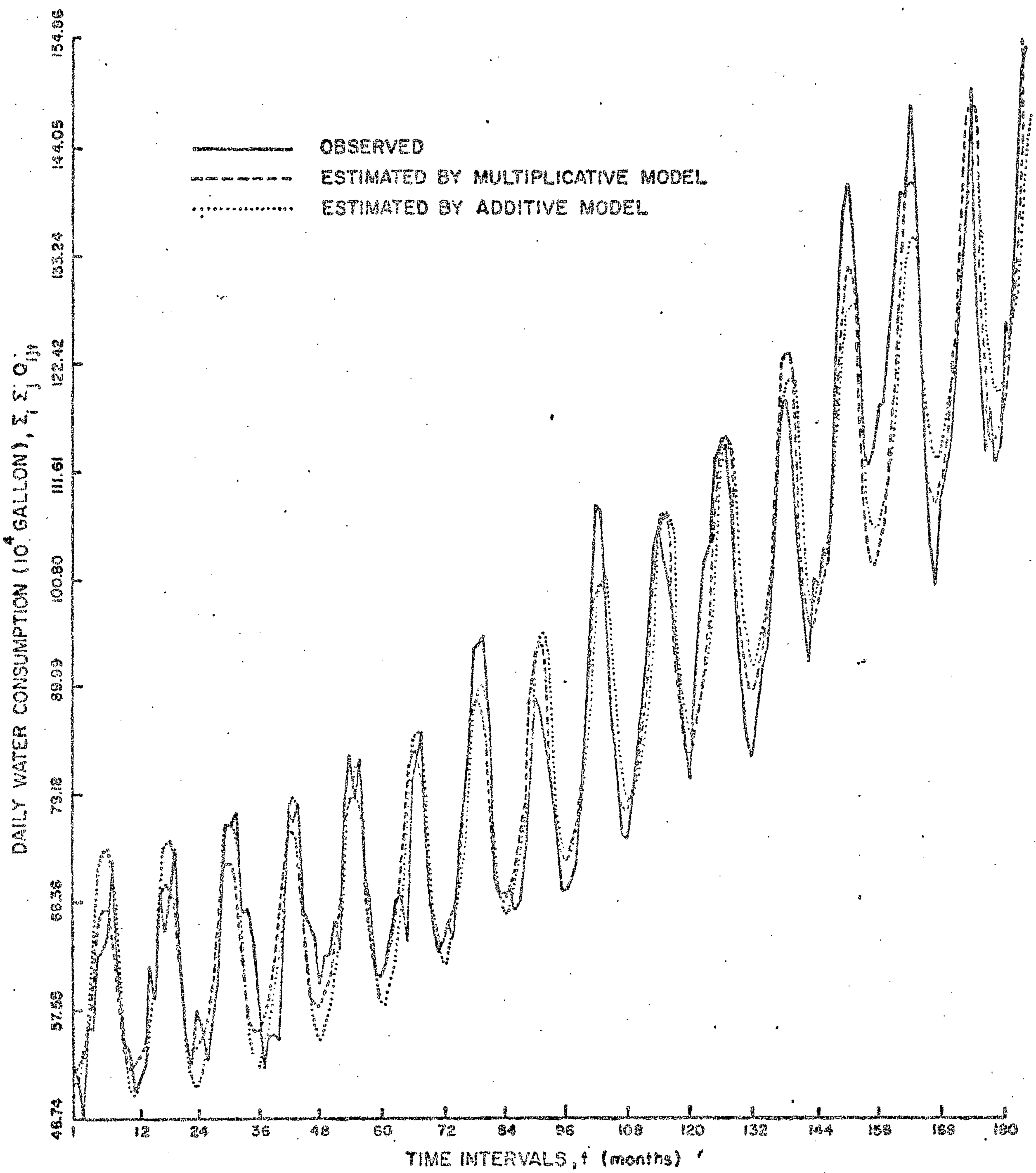

FIGURE 1. OBSERVED AND ESTIMATED BWS WATER CONSUMPTION FROM JANUARY 1961 TO JULY 1975

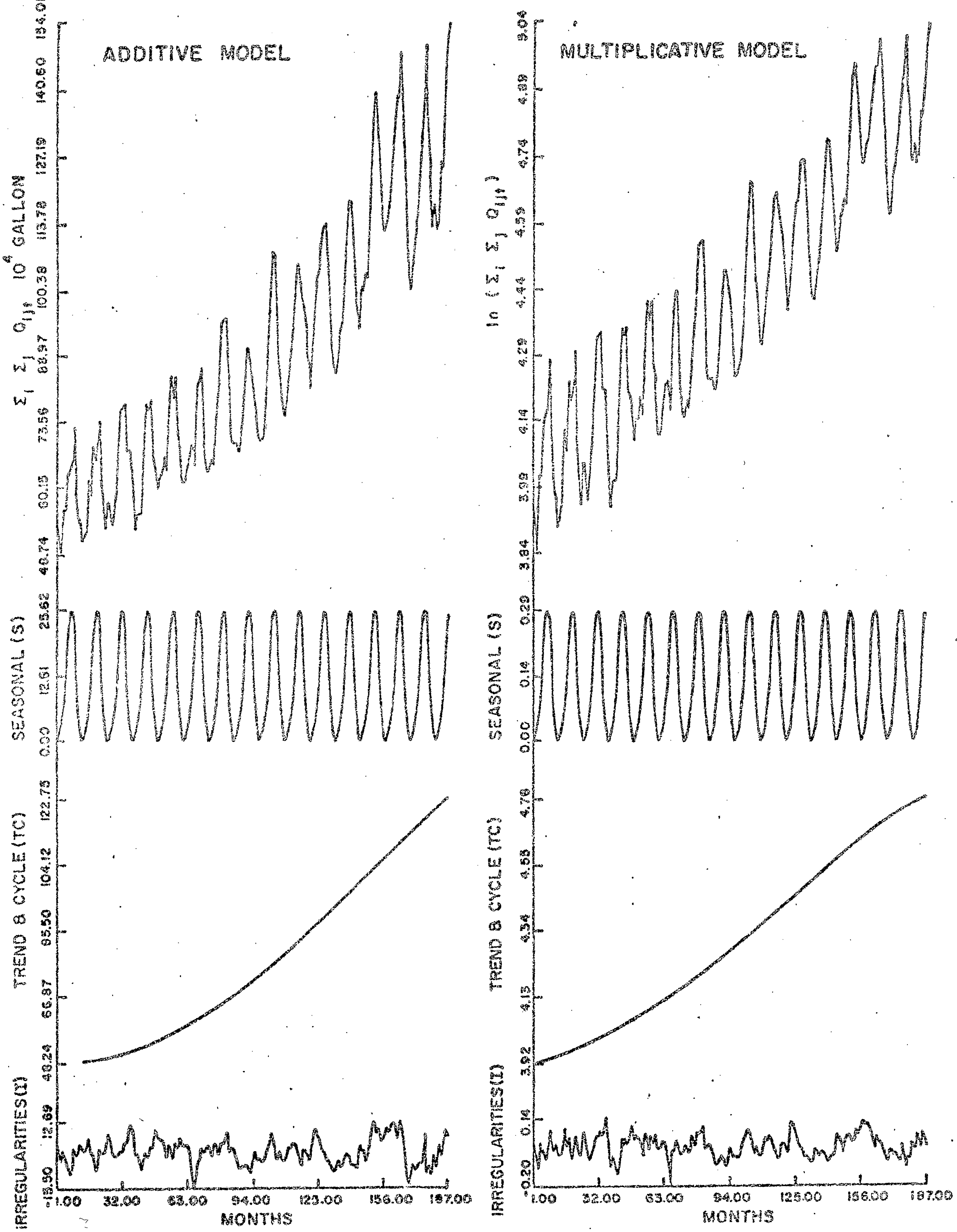

FIGURE 2. DECOMPOSITION OF BWS TOTAL WATER CONSUMPTION DATA (1961-1975)

cative model, the results are in log values. Therefore, while it is possible to simply add component parts in the additive model, in the multiplicative model antilogs must be computed from the summed log components.

Also closer inspection of the different components reveal interesting results. The levels of irregularities (I) in the two models tend to move in opposite directions as t increases. In the additive model, irregularities tend to increase with time, whereas in the multiplicative model the irregularities tend to decrease with time. This is consistent with the opposite time related tendencies of the two models.

The combined TC components in both cases exhibit steadily rising trends. These appear to be similar in properties but in fact there are differences. Although shown here on noncomparable scales, these curves can be compared with reference to the standardized regression coefficients reported in Tables 1 and 2. In the additive model, the first order coefficient α_1 of the polynomial in t is absent. This, in effect, reflects the definite curvalinear trend. When TC is estimated by the multiplicative model, α_1 is included and is equal to 0.348 in standardized form (Table 2). This not only tends toward flattening the curve when plotted on the log scale but also captures another factor which influences the level of the TC component. Without this factor, the TC level would tend to be lower throughout a good portion of the period.

Also, in both models a gradual tapering effect begins to show up in the later intervals. This results from the negative standardized coefficients for α_3. The effect is more pronounced in the multiplicative model. From Tables 1 and 2, the standardized α_3 is -0.210 for the multiplicative model as compared to -0.176 for the additive model.

The seasonal (S) components in both models appear as uniform fluctuations over time. This does not necessarily imply, however, that the estimated water consumption patterns are also uniform. The estimated levels of water consumption are computed using both the TC and S components. It turns out that a combination of curvalinear TC and uniform S pattern will result in nonuniform seasonal fluctuations in water consumption. The uniform S component merely reflects the regression estimates for the eleven monthly coefficients β_m (February to December). These estimates are computed using January of each year as a base month. Thus, there is no seasonal coefficient for January and the consumption level for this month is accounted for only by the TC level. Since TC has a rising trend, each 12th interval beginning with the first interval, results in a higher base month. Therefore, even if only one eleven-month pattern of β_m is repeated in successive years, the rising base months applied to this constant pattern of β_m will tend toward larger seasonal fluctuations. Also the curvalinear properties of TC contribute further to the nonuniform seasonal fluctuations from year to year.

Again the main difference between the two models is the additive and multiplicative manner in which these components are applied in estimating the level of water consumption. In the multiplicative model, changes in the levels of the TC and S components reflect more complex interactive factors which jointly affect the aggregate level of water consumption.

IMPLICATIONS

The implications of this study extend to both short- and long-run considerations. For long-run forecasting and capital investments planning purposes, both models begin with sufficiently high R^2 values to be potentially useful. However, in any forecasting model, errors tend to increase with distance into the future, and in the case of our two models, the multiplicative model appears to offer an important advantage. Errors in future intervals tend to be smaller with the multiplicative model.

Further refinements and analyses of these time series components are particularly useful for understanding the nature of seasonal peak demands and methods for reducing overinvestments in excess capacities. If disaggregated data is used in the regression analysis, the results reflected in the TC and S components can be usefully employed to making improvements in "high-low" checks which are conducted routinely by local water supply agencies.

The separation of the irregular component (I) offers important opportunities for better understanding the patterns of water consumption. This residual component is now assumed free of trend, cyclical, and seasonal influences. A variety of time series techniques can be used to analyze these residuals. For instance, adaptive forecast functions can be developed and applied to this component for the purpose of improving on short-run, e.g., quarterly, forecasts of water consumption and revenue receipts (Parzen 1974; Shiskin 1958; Box and Jenkins 1970).

Understanding of the behavior of the various time series components of water consumption data is essential for more effective analysis of economic incentive effects of alternative policy measures and closer integration of water supply and demand management.

REFERENCES

1. Akaika, H. "Fitting Autoregressive Models for Prediction," *Annal of Inst. Statis. Math.*, Vol. 21, 1969.

2. Alexander, Stuart M. "Multipurpose Computer for the Honolulu Board of Water Supply," paper presented at the Fall Conference of the California Section, AWWA in Honolulu, Hawaii, October 1974, 22 p.

3. Box, George E.P. and G.M. Jenkins. *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day, Chapter 9, 1970.

4. Durbin, J. "Trend Elimination for the Purpose of Estimating Seasonal and Periodic Components of Time Series," *Proceedings of the Symposium on Time Series Analysis*, Brown University, Providence, 1962.

5. Fishman, George S. *Spectral Methods in Econometrics*, Cambridge: Harvard University Press, 1964.

6. Granger, G.W.F. *Spectral Methods in Economic Time Series*, New Jersey: Princeton University Press, 1964.

7. Hannan, E.J. "The Estimation of Seasonal Variation in Economic Time Series," *JASA*, pp. 31-44, March 1963.

8. Huang, Wen-yuan. "Optimal Use of a Hawaii Agricultural Reservoir System," Ph.D. Dissertation, University of Hawaii, December 1974.

9. Jenkins, Gwilyn M. and Donald G. Watts. *Spectral Analysis and Its Applications*, San Francisco: Holden-Day, 1969.

10. Johnston, J. *Econometric Methods*, 2d ed. New York: McGraw-Hill, 1972, pp. 177-91.

11. Lobb, Howard J. "Demand Rate Economic," *Jour. Amer. Water Works Assoc.*, May 1975, pp. 246-50.

12. Moncur, James E.T. "An Exploratory Application of Two Methods of Analyzing Water Use Time Series," UH-WRRC Technical Rpt. No. 91, 1975, 16 p.

13. Oh, Ho-sung and Hiroshi Yamauchi. "An Economic Analysis of the Patterns and Trends of Water Consumption with the Service Areas of the Honolulu Board of Water Supply," UH-WRRC Technical Rpt. No. 84, 1974. 97 p.

14. Parzen, Emanual. "Some Recent Advances in Time Series Modelling," Institute of Electrical and Electronic Engineers, Inc., *Trans. on Automatic Control*, Vol. AC-19, No. 6, December 1974.

15. Shiskin, Julies. "Decomposition of Economic Time Series," *Science*, Vol. 128, No. 3338, December 1958.

16. Suits, D.B. "Use of Dummy Variables in Regression," *Amer. Stat. Assoc. Jour.*, December 1957, pp. 548-51.