MUNICIPAL WATER DEMAND STUDY OF WESTERN SKÅNE, SWEDEN
- Background Analysis with Some Preliminary Results

Tsuyoshi Hashimoto
Lennart de Maré

May 1980
WP-80-76

PREFACE

Water resource systems have been an important part of resources and environment related research at IIASA since its inception. As demands for water increase relative to supply, the intensity and efficiency of water resources management must be developed further. This in turn requires an increase in the degree of detail and sophistication of the analysis, including economic, social and environmental evaluation of water resources development alternatives aided by application of mathematical modelling techniques, to generate inputs for planning, design, and operational decisions.

During the year of 1978 it was decided that parallel to the continuation of demand studies, an attempt would be made to integrate the results of our studies on water demands with water supply considerations. This new task was named "Regional Water Management" (Task 1, Resources and Environment Area).

One of the case studies in this Task, carried out in collaboration with the Swedish Environmental Protection Board and the University of Lund, is the region of Western Skåne, Sweden. Although the Task emphasizes demand-resource integration, some analysis of the demand *per se* is necessary for better understanding of demand-generating factors. In this paper, background analysis concerning the municipal water demand characteristics of and the data base available for the region is presented with some preliminary results.

Janusz Kindler
Task Leader

ACKNOWLEDGEMENTS

We are especially grateful to Janusz Kindler, who put us on the right track, provided us with useful comments and gave a better structure to the paper.

We express our gratitude to Professor Z. Pawlowski, who provided us with many stimulating and encouraging suggestions at the initiation of this study.

Special thanks are due to Susan Arthur and Donald Erlenkotter for their extensive comments on the manuscript, which helped very much in improving this paper. Remaining deficiencies are, of course, our own responsibility.

This work was partially supported with funds provided by Stiftung Volkswagenwerk.

CONTENTS

MUNICIPAL WATER DEMAND STUDY OF WESTERN SKÅNE, SWEDEN - Background Analysis with some Preliminary Results

T. Hashimoto, and L. de Maré

I. INTRODUCTION*

This water demand study concerns the geographical region of the Malmöhus County, situated in South Sweden with an area of almost 5,000 km^2. Approximately 70% of the area is occupied by agricultural land. The population is presently about 740,000 i.e., the population density is 150 inhabitants per km^2 which is to be compared with the country average of 20 inhabitants per km^2. Malmöhus County is thus highly urbanized compared to the rest of the country.

Table 1 shows some land use and economic characteristics of the county as compared to the whole of Sweden. It is clearly indicated that the region has a potential for large conflicts between urbanization-industrialization and agricultural development. At the same time, the county's share of the potentially available water resources is in the order of only 0.5% with a pronounced lack of surface water resources.

The basic decision-making unit concerning the use of land, water, and other natural resources, is the municipality. After

---

*For more details about recent developments and the present situation in the region, see Andersson et al., (1979).

Table 1. Malmöhus County, percent shares of country totals.

Total Land			1.2
Population			9.0
Agricultural Land			10.4
Total Income			9.2
Manufacturing Industry, Value Added			9.1
:	Food and Beverages	20.8	
	Chemicals, Rubber etc.	18.3	
	Non-metallic Mineral Products	15.3	
	Textiles, etc.	8.7	
	Paper, Printing	8.0	
	Metal Products, Machinery, etc.	7.3	
	Basic Metal Industry	2.9	
	Wood and Wood Products	1.6	
	Others	4.5	

the municipal reform in 1974, Malmöhus County consisted of 21 municipalities (See Figure 1). This number, however, was reduced to 20 in 1977 when Bara was incorporated with Svedala. There are thus 20 to a large extent independent administrative units which, by tradition, have taken the responsibility to provide water of sufficient quantity and quality to cover water needs of various activities within their own jurisdictional boundaries.

The first larger water supply system based on surface water, the Vomb system, serving the City of Malmö, became operational in 1948. Already at that time, concerns about the future water supply of the Malmö region were expressed. It seemed to be evident that neither the groundwater nor the surface water resources of the county were sufficient for the future development. Thus in 1960, the Government initiated investigations on how to solve the future water supply of southwestern Sweden. The final report (SOU 1965:8) contained the first official attempt in Sweden to forecast municipal and industrial water use.

For the municipal water use, the household and general purpose (public, commercial, water works and losses) components

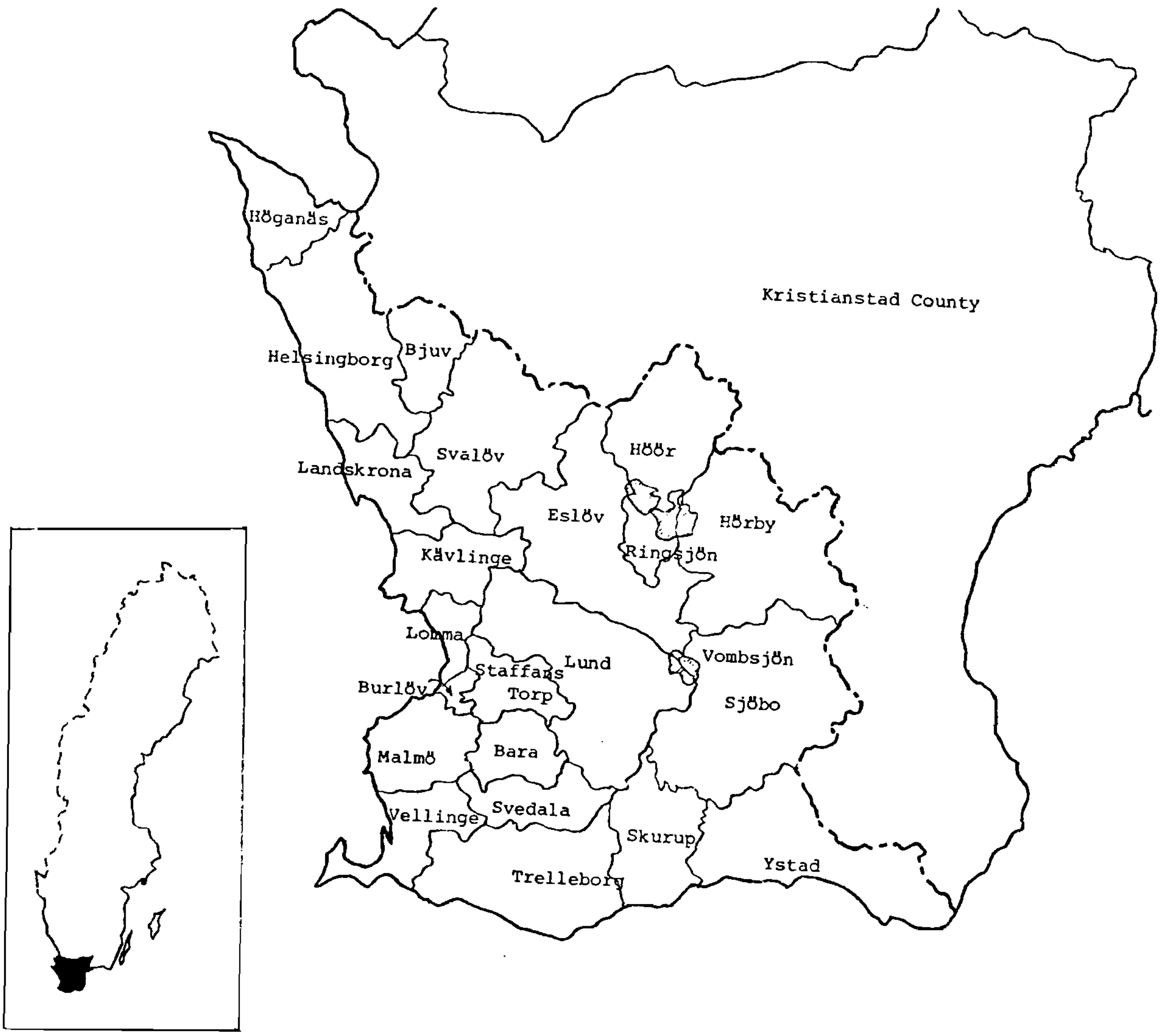


Figure 1. The study area--Malmöhus County, Sweden.

were forecasted based on official population forecasts and general per capita water use trends which were extrapolated to the years 1980 and 2000. It was thus estimated that the non-industrial urban per capita water use would increase from 200 liters per person and day in 1960 to 450 in the year 2000. For industrial water demand, it was considered impossible to adequately forecast the future situation. A rough estimate was given that it will double from 1960 to 1980 and treble until the year 2000.

In Figures 2 through 6, a comparison is made for the period 1967–1977 between the forecast and the actual statistics, as given by the Swedish Water and Sewage Works Association (VAV, S67–77). These data refer to 15 municipalities in western Malmöhus county and four neighboring municipalities in Kristianstad county. It is to be noted that the statistics refer to water supplied by the municipal water works. To the industrial water use statistics (and the total water use) for each year in Figure 2 should be added some 10 million m^3 per year to account for the self-supplied industry.

It is clear from Figure 2 that the total municipal water use has levelled off since 1970. Most of the deviation from the forecast falls on the non-industrial water use, while the industrial water use including the self-supplied water followed the forecast fairly well up to about 1974. An inconsistency in the population statistics caused by the municipal reforms in 1970 and 1974 is indicated by Figure 3. In spite of this, Figure 4 clearly shows that a large deviation of the per capita water use from the forecast has occurred. The reason for this deviation is, according to the Swedish Water and Sewage Works Association (VAV 1975), primarily that the household component of municipal water use is now close to a saturation value of about 220 liters per person and day.

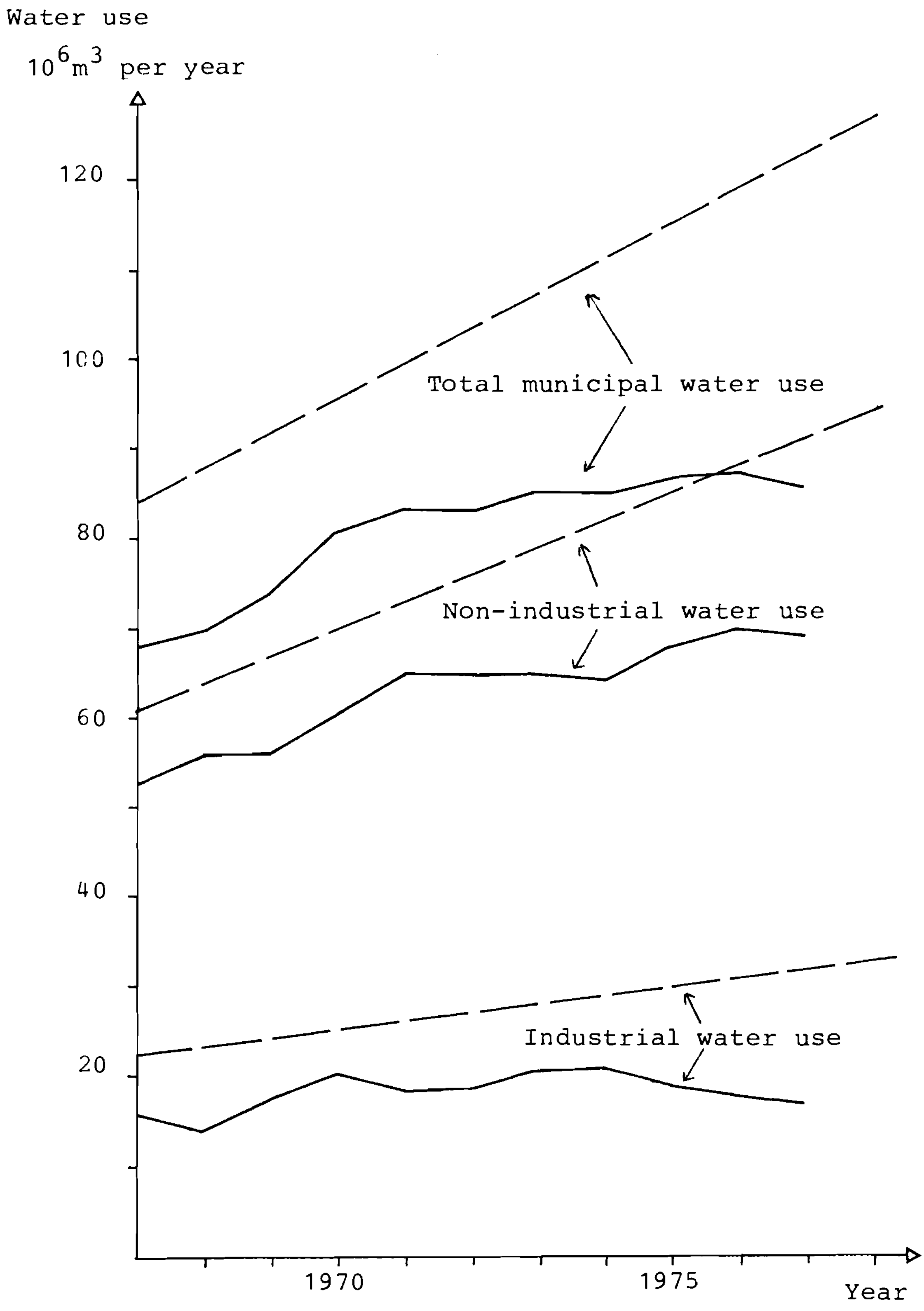


Figure 2. Municipal Water Use in the Region according to the 1965 Forecast (dashed lines) and the Observed Use (full lines).

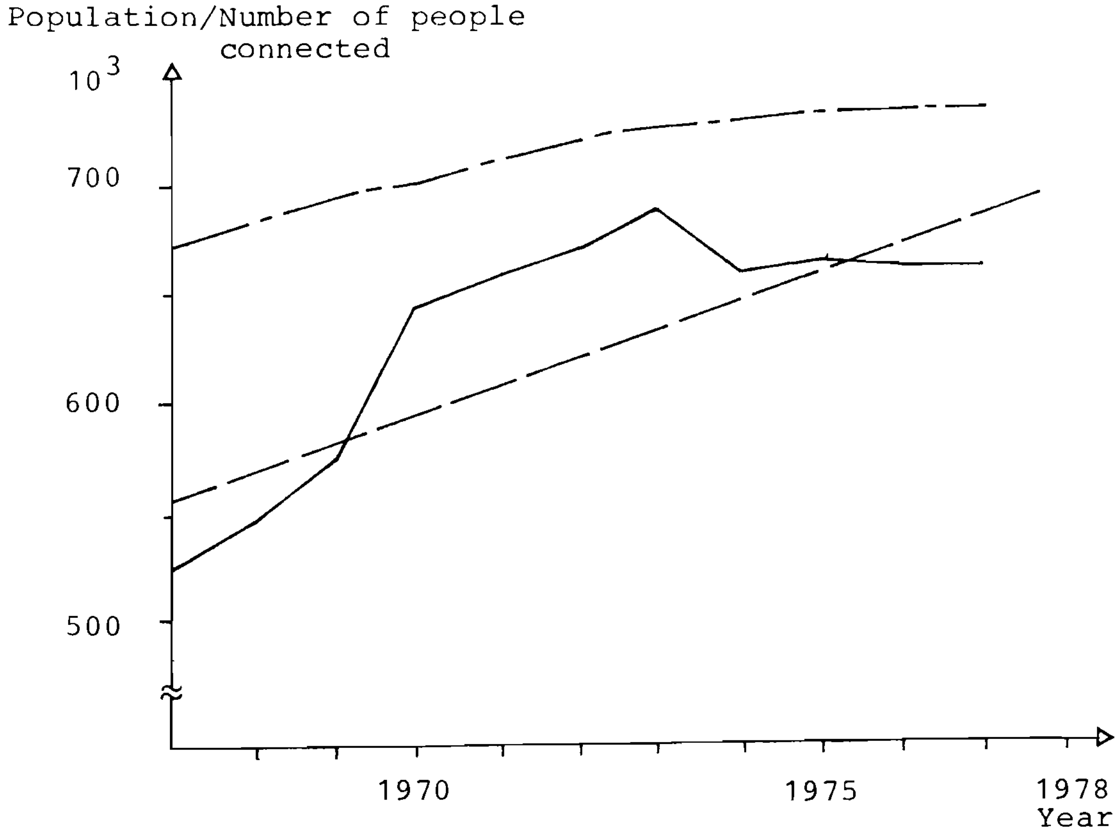


Figure 3. Number of people connected to the municipal water supply systems according to forecast (dashed lines) and statistics (full line). The upper curve shows total population.

Figure 4. Per capita water use according to forecast (dashed line) and statistics (full line).

If the regional statistics of Figures 2 through 4 are decomposed into subregions, one finds that the deviation from the forecast falls almost entirely on the most urbanized part of the county (Malmö, Lund, Staffanstorp, Svedala, Burlöv, Lomma), which accounts for 52% of the population. Other municipalities follow the forecasts relatively closely while for Eslöv-Kävlinge-Svalöv (8% of the population) the forecast for total municipal water use underestimated the actual use, see Figures 5 and 6. In all subregions, however, the per capita water use was considerably overestimated. This shift in water use from urban centers to surrounding municipalities is supposed to depend on the regional policy adopted in the early 1970's, leading to a decline in population growth for the whole region and especially for the urban centers. It is, however, of interest to see whether some of the statistical variables describing, for example, industrial activity and degree of urbanization can be related to the water use in the different municipalities.

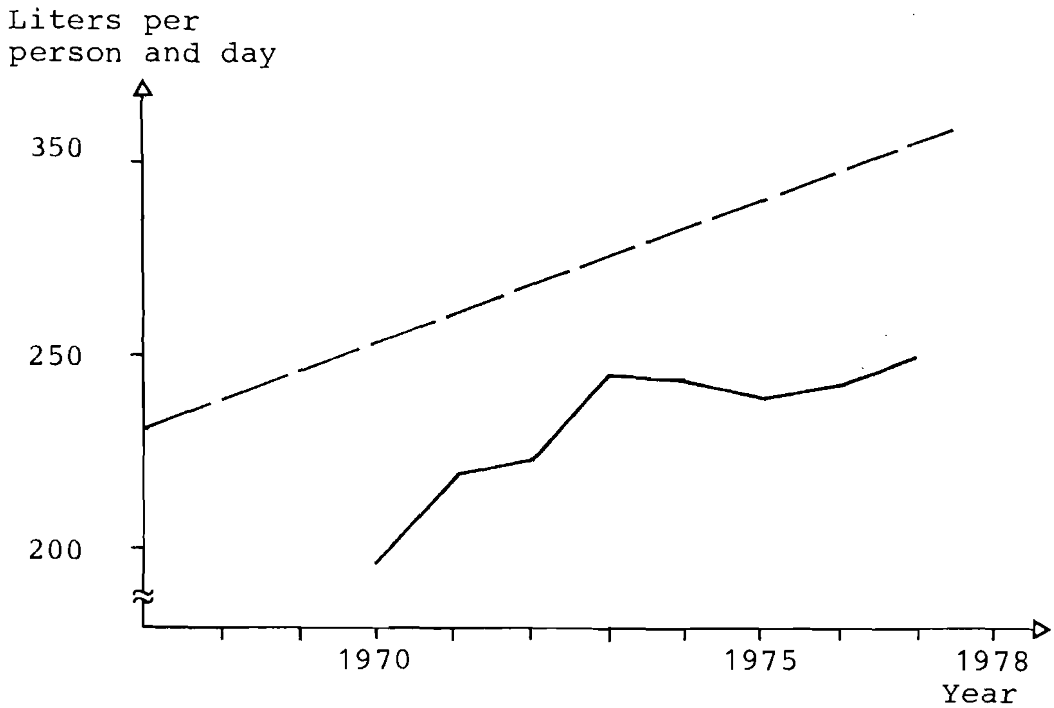
II. PURPOSES OF THE STUDY AND ITS ORGANIZATION

The primary purpose of this study is to show how water resources planners might go about modelling demand relationships concerning the municipal water use. Modelling of water demand relationships can assume a wide range of possible forms. This study employs the statistical approach which typically proceeds through the following steps: (1) assembly of the necessary data base, (2) identification of variables which affect the demand for water, (3) determination of the model structure, (4) estimation of model parameters, (5) model verification and validation,* and (6) practical utilization of the model so derived (Kindler et al., 1980).

---

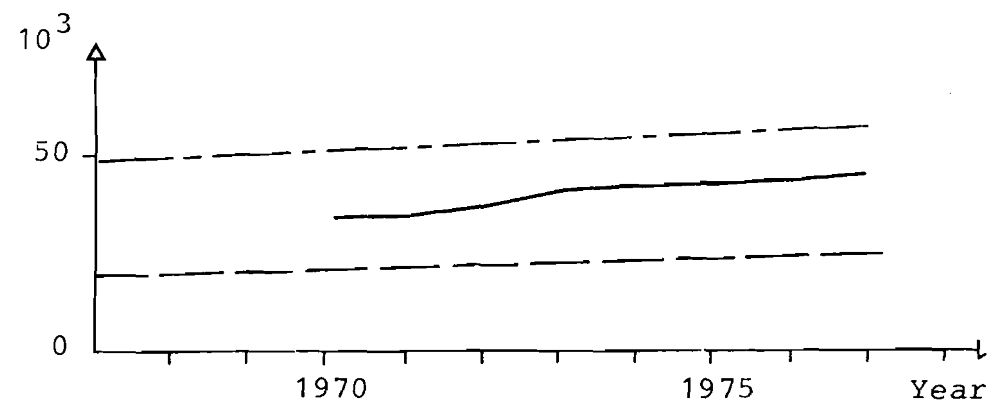
*Verification is understood as determining whether the "appropriate" model has been developed from a given <u>single</u> set of data describing water demand relationship. A <u>model</u> validation is the testing of the model's adequacy against an <u>independent</u> set of data.

Per capita water use



a) Per capita water use

Number of people connected



b) Number of people connected. The upper curve shows total population statistics.

Figure 5. Municipal Water Use in the Sub-region Eslöv-Kävlinge-Svalöv according to forecast (dashed lines) and statistics (full lines).

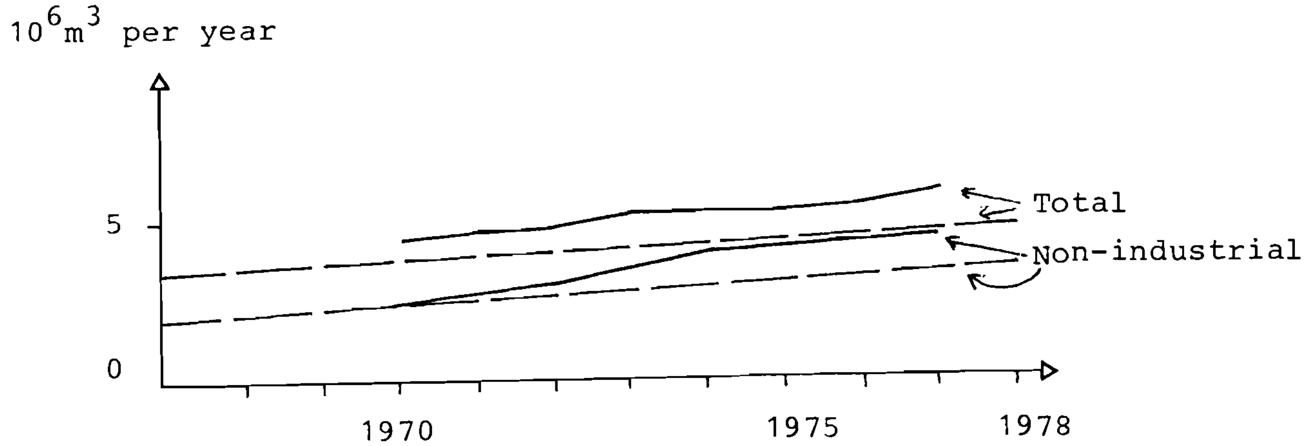


Figure 6. Municipal Water Use in the Sub-region Eslöv-Kävlinge-Svalöv according to forecast (dashed lines) and statistics (full lines).

Municipal water use covers a wide range of activities such as household operations, industrial production, commercial and public services--all receiving water through a common distribution system.

A major question in this municipal water demand study is how to take account of differences in activities among municipalities. Of twenty municipalities in the study area, some have more water-intensive industries connected to municipal supply systems; some have high recreational activities within their jurisdictional boundaries and many weekend houses which have quite different water use patterns than permanent residences. Also, wide variabilities are observed in the size of municipalities and their degree of urbanization.

Alternative approaches are possible to take account of the inter-municipality differences. Two principal questions have to be asked in choosing among these approaches: (1) Should one deal with each component--household, industrial, public, commercial, etc.--of the municipal water demand separately, or should they be aggregated into a total municipal water demand? (2) Should one use total gross or total per capita water demand as the dependent variable?* The answer to the first question depends on several factors, the most important being the availability of data on dependent and potential explanatory variables. Because of the character of available data on municipal water use in the Malmö county, the decision was made to model total (aggregated) demand with component differences represented by the appropriately selected explanatory variables. The second question remains to be answered by comparison of the results of demand analysis carried out for both types of dependent variables.

The following part of the paper is organized into five sections and follows generally the procedural steps presented above. The available data base is first surveyed in Section III to see what kind of analysis it allows. In Section IV, the explanatory variables that are considered of potential value in explaining the total municipal water demand are specifically identified. The procedures to develop better understanding of demand generating factors and, eventually to identify more important explanatory variables are described.

In Sections V and VI, the total gross and the total per capita municipal water demand are respectively used as dependent variables in the regression with appropriate model structures. Based on the results obtained and reported in these two sections, Section VII contains a discussion about the problem of forecasting, and Section VIII offers suggestions for extension of the study that is possible given the available data base; other possible approaches are also suggested.

---

*The total gross municipal water demand means the total amount of water demanded in a particular municipality in a year, and the total per capita demand is the total gross demand divided by the total population connected to the municipal supply system.

III. THE AVAILABLE DATA BASE

A number of studies have been undertaken to identify the factors that influence municipal water demand. The variables most commonly used are price, per capita income and climatological variables (Sewell and McClellan, 1979). These data are easily available for Swedish municipalities. The charges for water and sewage are published each year by the Swedish Water and Sewage Association (VAV, Tx69-77). The Swedish Central Bureau of Statistics (SCB) publishes each year the taxed income of the inhabitants in each municipality (SCB 67-77). Climatological variables have been disregarded in this study as the study area ($\sim 80 \times 60\ km^2$) is regarded as too small to show significant variations in this respect.

Other variables commonly used in identifying residential water demand are number of persons per dwelling or household, number of rooms per dwelling, age structure of household, property value, and urban area of municipality (Morgan, 1973; Darr, et al., 1975; Grunewald, et al., 1976; Gibbs, 1978). Statistics related to household characteristics are published every five years by the SCB in connection with the comprehensive population and housing accounts based on questionnaires (FOB 1970, 1975). These variables are thus available only for the accounting years. This is the same for the urban area variable.* Taxed property value is published the same way as taxed income. Age structure of household is not easily available. In this study, age of house has been used instead, complemented with a quality group classification.

In taking the industrial water use into account, a natural variable would be industrial production as used, for example, by Turnovsky (1969). The Swedish industrial statistics are extensive, especially for the mining and manufacturing industry, and are published each year by the SCB. However, difficulties

---

*Urban area is defined as an area with at least 200 inhabitants, where the distances between houses normally do not exceed 200 meters.

arise when decomposed statistics are needed: for example, sales value per municipality per industrial sector--information which usually requires special permission. For this study, the Malmöhus County Board provided the number of employees per sector and municipality. A comparison between the value added and the number of employees per sector for the whole county showed an almost linear relationship which justifies the use of the employment variable.

Other variables which are taken into account in this study are change in the price of water (recently used in a study by Sonnen and Evenson (1979)) and number of weekend houses. The latter variable is available from inventories made every five years by the County Board (MCB, 1977).

The water use statistics are published each year by the Swedish Water and Sewage Works Association (VAV, S67-77). Since 1970, these statistics are given per municipality. Total water use as well as its decomposition into household, general purpose, industrial, the water work's own use, and losses are given. However, there are a number of inconsistencies in the statistics, the most important being differences that exist between the municipalities in the definition and measurement of the components of water use. Moreover, the possibilities of basing the study on time series are limited because of the municipal reform in 1974; before 1970 the statistics were given per water work instead of per municipality, which makes comparisons almost impossible. It was thus decided to model total municipal water use (the most reliable statistics) using cross section data for 20 municipalities for years 1970 and 1975.

IV. ASSEMBLY OF DATA BASE AND DETERMINATION OF MODEL STRUCTURE

Enumeration of Potential Explanatory Variables

As the first step in the analysis of municipal water demand, all the variables which are considered of potential value in explaining the total municipal water demand are enumerated. As stated before, various activities in a municipality contribute to the observed overall water use pattern of the municipality. Each of these activities has to be represented by appropriate explanatory variables.

Data for the years 1970 and 1975 are available for the variables listed below which are preliminarily identified in the previous section as related to different components of the total municipal water demand. The following variables are mainly related to the household component:

1. Number of persons per household;
2. Number of households in one- and two-family houses;
3. Number of households in multi-family houses;
4. Number of apartments built in the period 1966-75;
5. Number of apartments in quality group 1-2.*

The industrial component can be represented by the following variables:

6. Value added for mineral product and manufacturing industry;
7. Sales value for mineral product and manufacturing industry;
8. Number of people employed in different sectors of industry.

Variables related to size of municipality and commercial, public and general uses category are:

9. Total population connected to municipal supply system;
10. Urbanized area;
11. Urban population.

The effects of recreational activities on the total municipal water demands can be represented by:

12. Number of weekend houses.

The following two variables are included in the data as related

---

* Apartments in quality groups 1-2 are supplied with water and sewage and contain WC, central heating and bath or shower room. Other quality groups lack one or more of these installations.

to price:

13. Variable costs charged to different consumers for water and sewage disposal;

14. Change in price of water and sewage from the previous year.

Finally, the following variables may be used to represent general standard of living which affects the water use behavior:

15. Annual taxable income of population connected;

16. Taxable value of non-agricultural property.

The property value is usually used as a surrogate of income, when data on the latter are not available. In this study, both property value and income are considered to see which one explains better water use behavior.

Preliminary Screening by Inspection

The analyst's common sense and judgement help to screen out some variables listed above by inspection. Variables 2 and 3 are complementary to each other with respect to the total number of households, therefore, variable 3 was eliminated. Variable 4 is screened out, since it is almost equivalent to variable 5.

Three variables are included in the data base to represent industrial activities. Since no sector-wise data are available on production statistics, variables 6 and 7 are eliminated. The number of people employed in water-intensive industry is used instead, where water-intensive industry includes the food and beverage, textile and clothing, chemical and petroleum (rubber and plastics sectors excluded), and metal goods (machine and transportation sectors excluded) sectors.

Variable 11 is practically equivalent to variable 9. This variable, however, is retained together with variable 10 in the data base, since there exists no better proxy to represent commercial, public and general uses component; variable 11 is transformed, however, into the ratio of urban population to the total population of the municipality.

Discriminating pricing* is used in some municipalities, but detailed data on the pricing structures are not readily available. Thus, the variable costs charged to the household users are used in this study. The price change is expressed in terms of costs charged to users for water and sewer in the current year, as a percentage of user costs in the previous year. This variable is included to see short-term response of water users to the price.

Data on the twelve potential explanatory variables preliminarily identified above are given in Table 2 for the year 1975. (This is the most recent year for which extensive data are available.)

Criteria for Selecting Appropriate Explanatory Variables

From the data base assembled in the previous subsection, the appropriate explanatory variables must be identified for each dependent variable. Analysis on real-world implications of explanatory variables helps to hypothesize which variables may affect the total municipal water demand, in what way and how significantly.

Statistically, there are several properties of explanatory variables to be considered in this screening process. The following two properties are generally accepted as important for explanatory variables:

i) high correlation with the dependent variable;

ii) low correlation with other explanatory variables.

When there are multiple candidates which have similar characteristics and satisfy these properties in the same degree, an additional property to be considered is:

iii) high relative variation in the observed values of the variable.

The relative variation can be measured by the coefficient of variation defined as the ratio between the standard deviation and the mean.

---

*In some municipalities in the study area, large industrial water users connected to municipal water supply systems get discount prices (i.e. lower rates for water they use).

Table 2. Data on explanatory and dependent variables for the year 1975.

Municipality[a)]	Explanatory Variable[b)]												Dependent Variable[c)]	
	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$	$\tilde{X}_9$	$\tilde{X}_{10}$	$\tilde{X}_{11}$	$\tilde{X}_{12}$	y	Y
Svalöv	7.4	6.49	137	220	329	2.53	399	1860	2616	61	110	330	358	970
Bara	5.0	1.94	87	210	199	3.35	1	997	1093	40	120	90	201	367
Staffanstorp	13.7	6.21	275	200	695	3.12	652	3081	3968	82	133	18	359	1700
Burlöv	13.8	5.00	299	140	630	2.60	1487	1723	5037	98	100	3	343	1730
Vellinge	17.5	15.40	360	120	1374	2.98	134	4878	5663	82	100	3303	238	1523
Bjuv	13.0	10.31	250	250	622	2.65	2026	2994	4241	91	125	57	767	3651
Kävlinge	16.2	10.18	316	140	1075	2.72	1339	3742	5264	80	127	947	321	1908
Svedala	6.4	3.16	123	250	274	2.50	70	1122	2347	75	167	129	268	630
Skurup	10.5	6.02	177	200	342	2.44	272	1864	2710	64	125	486	287	1098
Sjöbo	7.3	6.72	128	150	405	2.41	212	1877	2409	49	100	2066	258	698
Hörby	5.6	4.02	99	110	324	2.29	72	1354	2039	46	110	1402	261	530
Höör	7.4	7.55	139	120	417	2.42	334	1680	2285	72	120	2241	222	600
Malmö	235.0	69.67	5676	210	11084	2.07	12386	15825	114690	99	120	2753	353	30330
Lund	71.3	28.99	1655	218	3014	2.24	662	8172	30096	92	183	811	347	9068
Landskrona	33.0	15.19	718	295	1698	2.26	2348	3997	13615	91	118	1376	363	4390
Helsingborg	96.5	41.97	2130	240	4243	2.23	5455	10057	38415	94	141	1196	435	15361
Höganäs	18.3	12.19	357	249	1118	2.58	271	4852	6231	85	106	1483	250	1668
Eslöv	19.9	12.73	395	230	901	2.48	2036	3623	7125	72	100	336	390	2846
Ystad	20.5	10.21	412	250	797	2.25	870	3056	7021	74	137	1833	354	2655
Trelleborg	29.5	11.96	607	225	1293	2.41	1163	4168	9951	76	125	1678	302	3265

Table 2.: Notes:

a): The municipality of Lomma is not included, since the complete data are not available. Bara and Svedala, which are now a single municipality, are represented here separately.

b): $\tilde{x}_1$ = population connected to municipal water supply system in 10^3 persons

$\tilde{x}_2$ = urbanized area in km^2

$\tilde{x}_3$ = annual taxable income of population connected in 10^6 SKr

$\tilde{x}_4$ = variable costs charged to the users for water and sewer in 0.01 SKr/m^3

$\tilde{x}_5$ = taxable value of non-agricultural property in 10^6 SKr

$\tilde{x}_6$ = number of persons per household in urbanized area

$\tilde{x}_7$ = number of people employed in water intensive industry

$\tilde{x}_8$ = number of households in one- and two-family houses in urbanized area

$\tilde{x}_9$ = number of apartments in quality group 1-2 in urbanized area

$\tilde{x}_{10}$ = ratio urban population to total population in %

$\tilde{x}_{11}$ = price change (1975 price of water and sewer as % of 1974 price)

$\tilde{x}_{12}$ = number of weekend houses

c): y = total per capita municipal water use (yearly average) in ℓ/person/day

Y = total gross municipal water use in $10^3 m^3$/year

If the model is to be used for forecasting, another important property of explanatory variables is:

iv) ease of reliable forecasting.

These properties can be used as criteria for selecting appropriate explanatory variables at different stages of the analysis.

Determination of Model Structure

The selection of appropriate explanatory variables is closely connected with determination of model structure for each dependent variable. Both linear and non-linear forms are considered in this study--the effects of explanatory variables on the dependent variable are assumed to be either additive or multiplicative.

When total gross municipal water demand is used as the dependent variable in multiple regression, the population is clearly one of important explanatory variables. If the linear model is to be used to regress the total gross municipal water demand, each explanatory variable should also be expressed in gross terms--e.g. total income rather than per capita income (otherwise, the dimension is inconsistent). Then the explanatory variables are likely to be highly correlated with each other*; this is against the property (ii) listed above. The effects of population or size of municipality should better be separated from effects of other factors. Thus the log-linear is used in this study as a model structure for the total gross municipal water demand.

If the total per capita municipal water demand is regressed, the population effects are separated out from the analysis. In this case, the linear model may be used, based on the assumption that the effects of various water-using activities are additive to constitute the total per capita demand. The data base will be transformed so that the explanatory variables representing these activities have consistent dimension with the dependent variable.

---

*An example may illustrate this point. Suppose two explanatory variables x_1 (say, population) and X_2 (per capita income) have very low correlation. If relative variation of observed values (assumed all positive) of X_2 is much smaller than that of X_1, then X_1 and X_1X_2 (total income) are likely to be highly correlated, the effect of large variation in X_1 dominating.

Factor Analysis as a Guide for Selecting Explanatory Variables

The screening process to identify appropriate sets of explanatory variables can be aided by the use of factor analysis. Factor analysis is a technique to summarize the information contained in the data base for easier interpretation in the form of orthogonal factors (uncorrelated to each other) which are constructed as linear combinations of underlying variables. It allows clarification of the structure of the data base and identification of variables strongly associated with each factor.

There are several possible ways to use factor analysis in water demand study (McCuen et al., 1979; Sewell and McClellan, 1979), but empirical studies are still limited (Gum and Martin, 1977; Kim and McCuen, 1979).* Typically when there are a large number of explanatory variables highly correlated with each other and thus some variables may be redundant, factor analysis can be effectively used to reduce the number of variables. First the most significant factors themselves may be used as composite explanatory variables in regression--principal components regression (McCuen et al., 1979). Interpretation and measurement of these factors, however, may not be easy and their use in forecasting models may be disputable. Secondly, those variables strongly associated with the most significant factors may be selected as an appropriate set of explanatory variables (Kim and McCuen, 1979).

Use of factor analysis in water demand study can also be more incidental. It may happen, for instance, that a variable considered to represent a particular water use component has high correlation with the factor which appears to be the least associated with this component, judged from other variables having high correlation with this factor. This may suggest that the data base should be modified so that this water use component will be represented in a better way.

---

*The following discussions are mostly confined to uses of factor analysis for water demand study. For more general discussions on uses of factor analysis, see, for example Green (1978) or Bennett and Bowers (1976).

The utility of factor analysis in water demand study is supplementary; it can not be an exclusive tool for selecting explanatory variables. Other properties--e.g. correlation between explanatory variables and the dependent variables--should always be carefully checked.

Factor analysis is performed on the data base already assembled and given in Table 2. Elements of the factor matrix in Table 3, called factor loadings, are correlation coefficients between the variables and the factors. Each factor can be interpreted by looking at those variables having high (positive or negative) loadings in the respective factor.

The results in Table 3 show that the first factor accounts for about 62% of all the information contained in the data base, and many variables are strongly associated with this factor (i.e. have high factor loadings). This factor is strongly related to the size of municipality as seen from the high loadings associated with population and urbanized area (variables $\tilde{X}_1$ and $\tilde{X}_2$). High loadings of other variables in this factor are probably due to the fact that these variables are expressed in gross (total) terms rather than in per capita or in ratio.

This size factor is considered important in explaining the total gross municipal water demand, but may or may not be important for the total per capita demand. In either case, however, the size effects should be separated from the effects of other factors as stated before. This can be done by transforming those variables (except the population variable $\tilde{X}_1$ itself) having high factor loadings in the first factor into per capita values or other appropriate ratios. For instance, the urbanized area is expressed in ratio to the total area of municipality; the taxable income and property value are expressed in per capita values. The revised data are given in Table 4.

Factor analysis is again performed on the revised data base (see Table 4) and the results are given in Table 5. Four factors are identified as relatively significant. The first factor in Table 5 may be interpreted as representing the level of urbanization--the number of people connected to the municipal supply system, the ratio of urbanized to total area, the annual

Table 3. Factor matrix 1 (corresponding to the original data base given in Table 2)

Explanatory Variable	Principal factor[a)]	
	1	2
$\tilde{X}_1$	0.985	0.060
$\tilde{X}_2$	0.984	0.070
$\tilde{X}_3$	0.908	0.001
$\tilde{X}_4$	0.134	0.796
$\tilde{X}_5$	0.989	0.012
$\tilde{X}_6$	-0.524	-0.137
$\tilde{X}_7$	0.944	0.030
$\tilde{X}_8$	0.967	0.100
$\tilde{X}_9$	0.979	0.033
$\tilde{X}_{10}$	0.562	0.409
$\tilde{X}_{11}$	0.142	0.724
$\tilde{X}_{12}$	0.482	-0.612
Cumulative percentage of eigenvalue	61.9	76.1

a): Only two factors have an eigenvalue greater than unity.

Table 4. Revised data base for the year 1975.

Municipality[a)]	Explanatory Variable[b)]											
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
Svalöv	7.4	1.67	8.97	220	21.37	2.53	32.1	62	86	61	110	330
Bara	5.0	1.98	7.49	210	18.69	3.35	0.1	87	95	77	120	88
Staffanstorp	13.7	5.75	9.05	200	22.39	3.12	42.9	75	96	82	133	18
Burlöv	13.8	29.41	9.82	140	24.84	2.60	106.6	32	94	98	100	3
Vellinge	17.5	10.92	8.41	120	25.94	2.98	6.5	81	93	82	100	3303
Bjuv	13.0	8.96	9.07	250	24.32	2.65	143.6	60	84	91	125	57
Kävlinge	16.2	6.61	8.96	140	29.80	2.72	72.1	65	91	80	125	947
Svedala	6.4	2.68	8.71	250	17.89	2.50	9.2	45	93	75	167	129
Skurup	10.5	3.10	8.03	200	15.10	2.44	23.9	60	86	64	125	486
Sjöba	7.3	1.36	8.10	150	26.33	2.41	15.3	66	84	49	100	2066
Hörby	5.6	0.95	8.62	110	38.10	2.29	6.1	57	85	46	110	1407
Höör	7.4	2.58	9.13	120	28.02	2.42	34.8	64	86	72	120	2241
Malmö	235.0	45.54	12.08	210	27.15	2.07	50.2	13	91	99	120	2753
Lund	71.3	6.76	11.62	218	24.26	2.24	8.8	25	89	92	183	811
Landskrona	33.0	10.93	10.37	295	26.54	2.26	61.7	25	83	91	118	1376
Helsingborg	96.5	12.24	10.53	240	23.27	2.23	53.4	23	87	94	141	1126
Höganäs	18.3	8.54	9.22	249	27.91	2.58	13.1	69	89	85	106	1483
Eslöv	19.9	3.04	9.88	230	25.10	2.48	77.9	44	85	72	100	336
Ystad	20.5	2.92	9.76	250	21.16	2.25	36.6	39	87	74	137	1833
Trelleborg	29.5	3.54	10.20	225	23.42	2.41	33.2	37	87	76	125	1675

Table 4.: Notes:

a): The municipality of Lomma is not included

b): $X_i \equiv \tilde{X}_i$ for i=1,4,6,10,12

X_1 = population connected to municipal water supply system in 10^3 persons

X_2 = ratio urbanized area to total area in %

X_3 = annual taxable income per capita in 10^3SKr/person

X_4 = variable costs charged to users for water and sewer in 0.01 SKr/m^3

X_5 = taxable non-agricultural property value per capita in 10^3 SKr/person

X_6 = number of persons per household in urbanized area

X_7 = number of people employed in water intensive industry per 1000 population

X_8 = ratio of households in one- and two-family houses to total number of households in %

X_9 = ratio apartments in quality group 1-2 to total number of apartments in %

X_{10} = ratio urban population to total population in %

X_{11} = price change (1975 price of water and sewer as % of 1974 price)

X_{12} = number of weekend houses

Table 5. Factor matrix 2 (corresponding to the data base given in Table 4)

Explanatory Variable	Principal factors[a)]			
	1	2	3	4
X_1	0.890	-0.118	-0.003	-0.085
X_2	0.871	0.205	-0.155	0.276
X_3	0.878	-0.266	0.172	-0.022
X_4	0.097	-0.369	0.761	0.155
X_5	0.145	-0.092	-0.842	0.054
X_6	-0.434	0.866	-0.001	0.104
X_7	0.204	-0.114	0.166	0.881
X_8	-0.719	0.552	-0.281	-0.039
X_9	0.199	0.901	0.075	-0.190
X_{10}	0.716	0.364	0.327	0.293
X_{11}	0.236	-0.006	0.626	-0.573
X_{12}	-0.363	-0.319	-0.762	-0.271
Cummulative percentage of eigenvalue	35.4	57.0	74.5	85.8

a): Only those factors having the eigenvalue greater than unity are shown.

taxable income per capita, and the ratio of urban to total population have high factor loadings. Variables X_6, X_8, and X_9, representing housing types have high loadings in the second factor, and fourth factor clearly represents the level of industrial activities. The third factor is more difficult to interpret. The price variables have relatively large positive values of factor loading, while the non-agricultural property value and the number of weekend houses have large negative values.

It is interesting to note that the income variable X_3 has high loading in the urbanization factor but not in the third factor which appears to be more directly related to water use behavior. Instead, the property value X_5 has large negative loading in this factor.

Any variable having a high factor loading for a particular factor can be used in regression to represent this factor.* In this way, a set of explanatory variables satisfying the property (ii) listed before - low correlation among explanatory variables - is likely to be identified, since the factors themselves are independent of each other. Of course, the correlation between the explanatory variables and the dependent variable should also be checked. When there are several variables with high factor loadings (as in the case with our data base) for a given factor and the same level of correlation with the dependent variable, other criteria such as high relative variation (the property (iii)) can be used to select among them.

V. TOTAL GROSS MUNICIPAL WATER DEMAND

In Section IV, twelve variables were identified as potentially important in explaining the total gross municipal water demand. A way of using factor analysis to select the most appropriate sets of explanatory variables from these twelve variables was also suggested.

Four principal factors were identified by factor analysis. These are called (i) the urbanization factor, (ii) the housing factor, (iii) the demand sensitivity factor, and (iv) the

---

*It is possible to choose more than one variable from each factor, but this may not be very meaningful. Besides, it may lead to irrational sign and/or magnitude of regression coefficients due to multicollinearity.

industrialization factor. Variables having high factor loadings in these factors constitute an appropriate set of explanatory variables.

Variable X_1 (population connected) is selected to represent the urbanization factor. Variable X_5 has the largest absolute value of factor loading in the demand sensitivity factor but its relative variation of observed values is much lower than variable X_{12}, which has the second largest factor loading. Variables X_6 and X_9 have almost the same factor loading on the housing factor. The industrialization factor is naturally represented by variable X_7.

In addition to the six variables identified above, the price variables (X_4 and X_{11}) are also considered in regression, since pricing is considered to be a major tool to affect water use, if the water demand is price elastic. Stepwise regression is run on these eight variables with the log-linear form assumed to be an appropriate model structure for the total gross municipal water demand.* The results are summarized in Table 6.

The results show that the population connected (variable X_1) is the most important one in explaining the total gross municipal water demand (as expected), and the industrialization factor (variable X_7) is the next important variable. The regression coefficient associated with the variable X_9 is apparently irrational, since according to the model the effect of this variable on the dependent variable is in one order of magnitude larger than the effects of any other variable.

The positive values of the regression coefficients associated with the price variables X_4 and X_{11} are not necessarily irrational *per se*, but their validity in water <u>demand</u> relationship is very questionable. The regression coefficients may not

---

*The stepwise multiple regression programme in the SSP Library (IBM System/360 Scientific Subroutine Package) was used for the present study. This programme selects the explanatory variables one by one in the order of their importance. The criterion of importance is based on the reduction of the sum of squares of observed values of the dependent variable.

Table 6. Stepwise regression of total gross municipal water demand.

Explanatory Variables \ Step	Regression coefficient[a)] (t-value[b)])							
	1	2	3	4	5	6	7	8
X_1 = population connected	1.112 (17.61)**	1.047 (19.14)**	1.007 (19.33)**	1.042 (19.05)**	1.115 (18.83)**	1.165 (17.70)**	1.179 (18.27)**	1.154 (17.01)**
X_4 = variable cost charged			0.377 (2.26)*	0.324 (1.98)+	0.183 (1.15)	0.162 (1.05)	0.077 (0.48)	0.208 (1.05)
X_5 = non-agricultural property value per capita								0.290 (1.12)
X_6 = number of persons per household						0.804 (1.49)+	1.115 (1.96)+	1.005 (1.76)+
X_7 = ratio employed in water-intensive industry		0.114 (3.27)**	0.115 (3.68)**	0.105 (3.42)**	0.057 (1.62)+	0.062 (1.84)+	0.066 (2.01)+	0.062 (1.90)+
X_9 = ratio appartments in quality group 1-2					-2.972 (2.19)*	-4.124 (2.72)*	-4.907 (3.12)**	-4.949 (2.99)*
X_{11} = price change 1974-75							0.366 (1.37)+	0.450 (1.64)+
X_{12} = number of weekend houses				-0.040 (1.55)+	-0.088 (2.77)*	-0.083 (2.69)*	-0.085 (2.84)*	-0.093 (3.06)*
intercept	84.44	71.38	12.63	19.71	3.01 $\times 10^7$	2.31 $\times 10^9$	15.13 $\times 10^9$	1.03 $\times 10^9$
R*	0.972	0.982	0.986	0.987	0.990	0.990	0.991	0.991
R^2	0.945	0.966	0.974	0.978	0.934	0.986	0.988	0.989
F	309.96	243.95	203.68	166.88	168.18	152.78	140.07	125.20

a): The first step column, for instance, reads as follows: $Y = 84.44\ X_1^{1.112}$

b): The double and single asterisks and the cross denote the significance at 1%, 5% and 20% levels respectively.

reflect the price response of, or the price effects on, the water users since the causal relationships may be rather the other way around (e.g. high demand ⟶ stress on supply ⟶ higher price).*

Excluding the price variables and variable X_9, and representing the demand sensitivity factor by either the variable X_5 or X_{12}, the following water use equations can be identified:

$$Y = 95.30\ X_1^{1.073}\ X_5^{-0.244}\ X_6^{0.410}\ X_7^{0.126}\ ,\quad R^2 = 0.970$$

(17.25**) (1.02) (0.76) (3.38**)   R* = 0.982, F=119.74

$$Y = 102.47\ X_1^{1.081}\ X_6^{-0.086}\ X_7^{0.098}\ X_{12}^{-0.054}\ ,\quad R^2 = 0.972$$

(18.10**) (0.14) (2.55*) (1.61+)   R* = 0.983, F=131.78

The double and the single asterisks and the cross after the t-value given in the parenthesis here and thereafter, denote the significant levels at 1%, 5% and 20%, respectively.

The value of the coefficient of determination, or the multiple correlation coefficient R squared is very high for either model. This, however, should not be taken as a sign of goodness-of-fit by itself. The residuals of the estimates are important criteria to judge goodness-of-fit, when the practical use of the model is of concern. The estimates of the second model above, range from 0.541 (for municipality of Bjuv) to 1.378 (for Burlöv) times the observed water use figures.

Variable X_6 representing the housing factor is relatively insignificant as seen from the small t-value and the inconsistent sign associated with its regression coefficient in either model. This is probably due to the fact that the household component of the total municipal water use does not vary much among municipalities. Eliminating this variable, the following alternative water use equations are obtained:

---

*General treatment of this phenomena leads to the so-called "identification problem" (see for example, Johnston, 1972). Given only limited observations on price and water demand, there is no *a priori* reason to believe that the observed price is negatively correlated with the water use, even if the water demand is elastic. See also Section VII.

$$Y = 152.78\ X_1^{1.052}\ X_5^{-0.245}\ X_7^{0.117} \quad , \quad R^2 = 0.968$$
$$(19.20^{**})\ (1.04)\ (3.36^{**}) \qquad R^* = 0.982,\ F=163.79$$

$$Y = 91.82\ X_1^{1.084}\ X_7^{0.101}\ X_{12}^{-0.051} \quad , \quad R^2 = 0.972$$
$$(19.76^{**})\ (3.04^{**})\ (1.86^{+}) \qquad R^* = 0.984,\ F=187.15$$

The regression coefficient associated with the population (variable X_1) is close to unity in any of the models above. This partly justifies treating the total gross municipal water demand as the per capita demand times the population. In the next section, the per capita municipal water demand is used as the dependent variable.

VI. TOTAL PER CAPITA MUNICIPAL WATER DEMAND

Before selecting the most appropriate explanatory variables for the total per capita municipal water demand, minor modifications must be made to the data base assembled in Section IV. The variable X_1 (population connected) is removed, and the number of weekend houses (variable X_{12}) is transformed into per capita value to make the dimension consistent with the dependent variable. This does not change the structure of the data base in any significant way.

Factor analysis of the modified data base identifies four principal factors (see Table 7). Again, these factors may be called (i) the urbanization factor, (ii) the housing factor, (iii) the demand sensitivity factor, and (iv) the industrialization factor. Roughly speaking factors (i), (ii) and (iv) may correspond to commercial/public, household and industrial components of the total per capita municipal water demand, respectively.

The procedure followed in Section V can again be used to model the total per capita municipal water demand. First, the variables having high loadings for a given factor are selected as candidates to represent that factor. The relative variation of each variable and its correlation with the dependent variable are also considered. This leads to use of the following variables:

Table 7: Factor Matrix 3

Explanatory variable[a]	Principal factor[b]			
	1	2	3	4
x_2	0.879	-0.127	-0.187	0.170
x_3	0.706	-0.613	0.078	-0.134
x_4	-0.001	-0.385	0.778	0.121
x_5	0.071	-0.114	-0.857	0.123
x_6	-0.035	0.972	0.048	0.040
x_7	0.335	-0.145	0.224	0.803
x_8	-0.469	0.810	-0.210	0.022
x_9	0.501	0.723	0.033	-0.343
x_{10}	0.884	0.036	0.272	0.081
x_{11}	0.173	-0.182	0.514	-0.694
x_{12}	-0.550	-0.750	-0.750	-0.083
Cumulative percentage of eigen-value	37.0	57.0	73.8	84.9

a: The variables correspond to the data base in Table 4 $x_i = X_i$ for i = 2,...,11) except the variable X_{12} has been transformed into per capita values x_{12}.

b: Only those factors having the eigenvalue greater than unity are shown.

for the urbanization factor: x_2, x_{10};
for the housing factor: x_6, x_8;
for the demand sensitivity factor: x_5, x_{12}; and
for the industrialization factor, x_7.

Assuming linear structure of the model, stepwise regression is used on the seven variables listed above, plus the price variables x_4 and x_{11}.

The results are summarized in Table 8. First the ratio of people employed in water-intensive industry (variable x_7) is found most significant in explaining the total per capita municipal water demand. Second, the price variables x_4 and x_{11} enter the regression with the positive coefficients as before.

The results of stepwise regression, run without the price variables, are given in Table 9. If the variable that enters the regression first is selected as the best explanatory variable for each factor, the following water <u>use</u> equation is obtained:

$$y = 427.68 - \underset{(1.44^{+})}{2.334\, x_2} - \underset{(1.26)}{63.51\, x_6} + \underset{(3.30*)}{2.524\, x_7} - \underset{(1.32)}{0.243\, x_{12}},$$

$$R^2 = 0.722,$$

$$R* = 0.820, \quad F = 10.40.$$

The values of regression coefficient associated with variable x_6 in Table 8 and Table 9 are very large, and introduction of this variable changes the intercept of the regression equation significantly. No reasonable justification can be given to explain this except that it is partly due to the low variation in observed values of this variable. If variable x_8 is used instead to represent the housing factor, the following alternative is obtained:

$$y = 294.60 - \underset{(1.35^{+})}{2.381\, x_2} + \underset{(5.21**)}{2.556\, x_7} - \underset{(0.77)}{0.672\, x_8} - \underset{(0.89)}{0.162\, x_{12}}$$

$$R^2 = 0.705$$

$$R* = 0.808, \quad F = 9.58.$$

Table 8: Stepwise Regression of the Total Per Capita Municipal Water Demand

Explanatory variables \ Step	Regression coefficient[a/] (+ - value [b/])								
	1	2	3	4	5	6	7	8	9
x_2 = ratio urbanized area							1.261 (0.68)	2.697 (1.02)	2.515 (0.89)
x_4 = variable cost charged			0.575 (1.49*)	0.720 (1.77*)	0.758 (1.96*)	0.984 (2.25*)	1.108 (2.30*)	1.275 (2.38*)	1.218 (2.05*)
x_5 = non-agricultural property value per capita						3.414 (1.09)	3.390 (1.06)	4.282 (1.24)	4.619 (1.22)
x_6 = number of persons per household					-120.03 (1.64*)	-104.26 (1.41*)	-132.40 (1.54*)	-89.51 (0.86)	-107.66 (0.85)
x_7 = ratio employed in water-intensive industry	2.610 (6.00**)	2.785 (7.00**)	2.610 (6.49**)	2.755 (6.52**)	2.966 (7.03**)	2.981 (7.10**)	2.968 (6.93**)	3.120 (6.53**)	3.071 (5.82**)
x_8 = ratio households in one- and two-family houses				0.846 (1.08)	2.573 (1.99*)	2.599 (2.03*)	3.412 (1.93*)	3.144 (1.71*)	3.327 (1.65*)
x_{10} = ratio urban population								-1.624 (0.77)	-1.624 (0.74)
x_{11} = price change 1974.75		1.670 (2.23*)	1.258 (1.62*)	1.519 (1.88*)	1.830 (2.31*)	2.147 (2.56*)	2.350 (2.59*)	2.873 (2.51*)	2.776 (2.23*)
x_{12} = number of weekend houses									-0-083 (0.28)
intercept	224.55	29.63	6.49	92.67	73.58	-119.75	-140.23	-229.65	-179.51
R*	0.813	0.849	0.860	0.862	0.876	0.878	0.874	0.870	0.860
R^2	0.661	0.734	0.765	0.781	0.814	0.829	0.835	0.842	0.843
F	36.98	24.87	18.45	14.26	13.15	11.29	9.37	8.02	6.59

a: The first step column, for instance, reads : $y = 224.55 + 2.610\ x_7$.

b: The double and single asterisks and the cross denote the significance at 1%, 5% and 20% levels, respectively.

Table 9. Stepwise Regression of the Total Per Capita Municipal Water Use (without price variables)

Explanatory Variables \ Step	Regression coefficient (t-value[a)])						
	1	2	3	4	5	6	7
x_2 = ratio urbanized area		-1.483 (0.95)	-1.839 (1.15)	-2.334 (1.44+)	-1.639 (0.90)	-2.622 (1.21)	-2.726 (1.19)
x_5 = non-agricultural property value per capita							0.882 (0.20)
x_6 = number of persons per household				-63.51 (1.26)	-161.24 (1.30)	-205.98 (1.53)+	-206.88 (1.48)+
x_7 = ratio employed in water-intensive industry	2.610 (6.08)**	2.754 (6.03)**	2.606 (5.43)**	2.524 (5.30)**	2.532 (5.27)**	2.469 (5.04)**	2.452 (4.77)**
x_8 = ratio households in one- and two-family houses					1.804 (0.87)	2.475 (1.11)	2.469 (1.06)
x_{10} = ratio urban population						1.703 (0.87	1.647 (0.80)
x_{12} = number of weekend houses			-0.179 (1.00)	-0.243 (1.32)	-0.387 (1.56)+	-0.333 (1.29)	-0.369 (1.15)
intercept	224.55	231.72	253.67	427.68	585.52	537.52	528.14
R*	0.813	0.812	0.813	0.820	0.818	0.816	0.801
R^2	0.661	0.677	0.695	0.722	0.735	0.749	0.750
F	36.98	18.84	12.89	10.40	8.34	6.96	5.56

a): The double and single asterisks and the cross denote the significance at 5%, 10% and 20% respectively.

Despite the lower value of the multiple correlation coefficient R^2 (or R*) for the model above, the fit is better than in the models of the total gross municipal water use, given in Section V, with respect to the residuals of the estimates. The relative residual of the model above takes the maximum value 35% for the municipality of Burlöv, but is less than 20% for most of the municipalities (see also Figure 7). The log-linear form of regression equation was also tested, but the fit in general was poorer.

VII PROBLEM OF FORECASTING

Regression analyses have been made on the cross-section data for twenty municipalities in Malmöhus County in 1975, by using the total gross and the total per capita municipal water demand as the dependent variables. The cross-section studies, however, are static in nature and their use for forecasting is not unconditionally justified as discussed in the following.

If all the municipalities have the same aggregated demand curve - i.e. the same price response, given all other demand-affecting factors (e.g. welfare level), and if effects of factors including price are properly represented by a set of explanatory variables, then straightforward regression on this set of explanatory variables using cross-section data will provide an appropriate water demand relationship. Note in this case that each municipality may have a different supply curve, but this does not pose any serious problem as illustrated by Figure 8(a). In this figure is illustrated that the intersections between the common demand curve and the different supply curves S_1, S_2 and S_3 allow inference of the relevant demand curve.

Such a fortunate situation as described above does not usually exist in reality for two major reasons. First, all the demand affecting factors may not be appropriately represented by the selected explanatory variables, even if all the municipalities have basically the same water demand relationship. Secondly, even if practically all other demand affecting factors are taken into account, price response of municipalities may be inherently different. The latter may very well be the case, since water

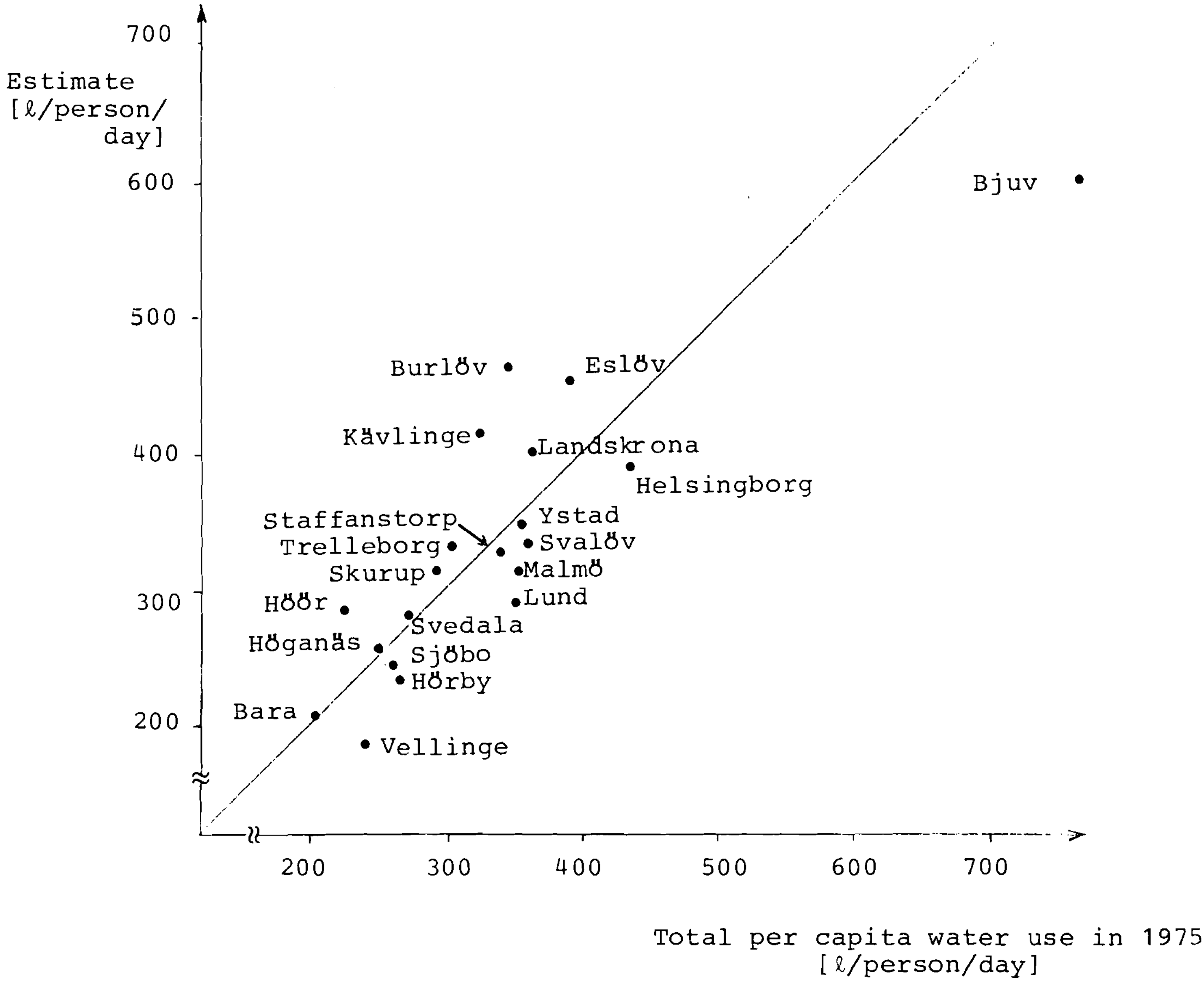


Figure 7. Comparison of the Estimated and Actual Per Capita Municipal Water Use in 1975

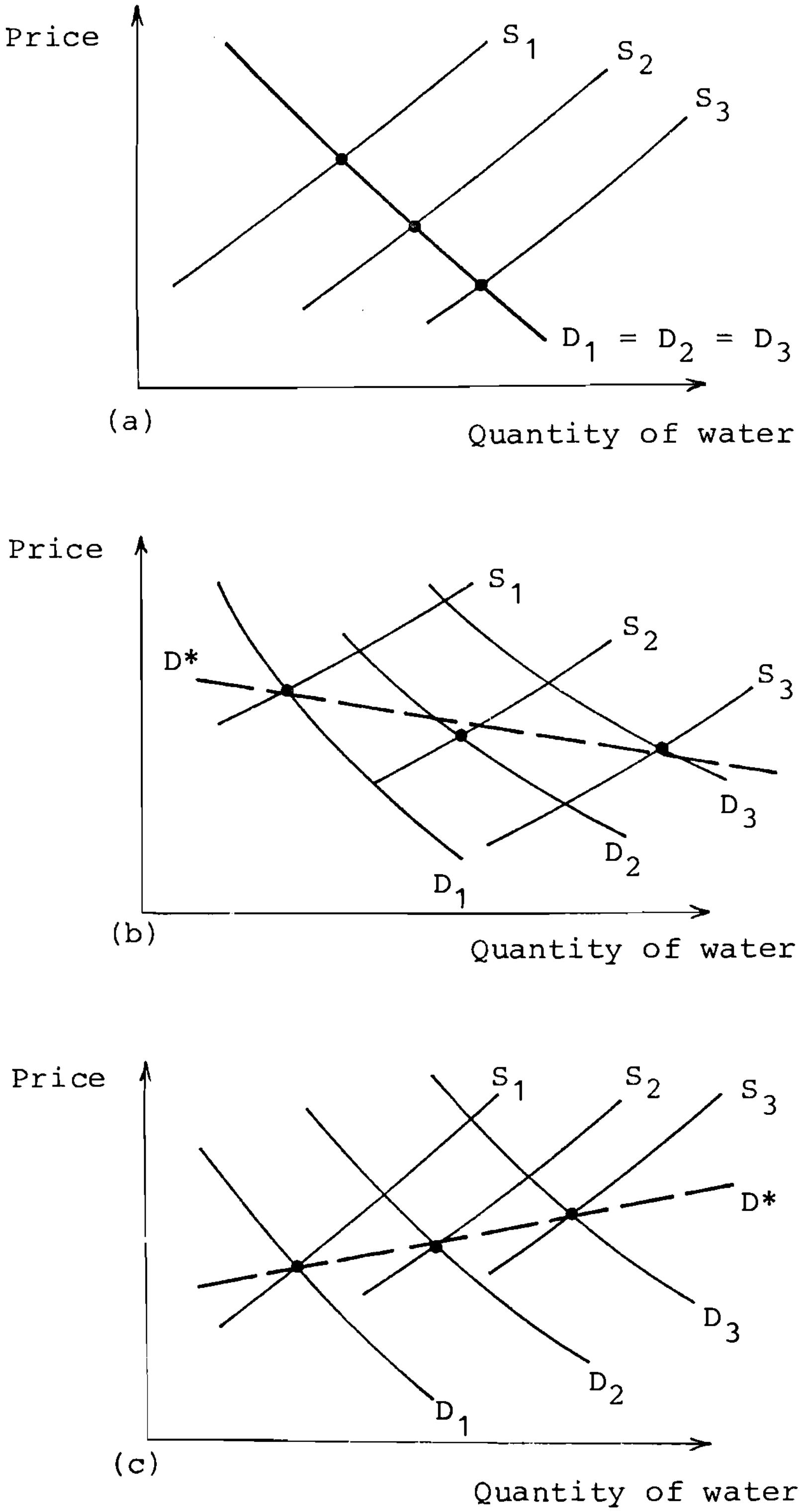


Figure 8. Possible Relationships of Aggregated Supply and Demand Functions S_1, S_2, S_3 and D_1, D_2, D_3 for Three Municipalities.

users aggregated into municipalities do not usually represent random sampling. In either case, municipalities have different demand curves, even if all other explanatory variables have the same values for all the municipalities. It is under this situation that differences on supply side, which certainly exist among municipalities, complicate identification of water demand relationship as illustrated by Figure 8(b). Knowledge of only three intersections between demand curves D_1, D_2 and D_3 and supply curves S_1, S_2 and S_3 may lead to inference of an erroneous demand curve D* in Figure 8(b).

The point discussed above can be more specifically illustrated referring to the results obtained in this study. It was found in both previous sections that the price variables entered the regression with positive coefficients. Clearly such a regression equation is invalid as the water demand relationship and thus cannot be used for forecasting purposes. The regression coefficients do not reflect the price response of or the price effects on the water users. The situation may be illustrated by Figure 8(c), where an upward sloping "demand curve" D* is inferred based on three observed demand-supply intersections.

How can this situation be remedied so that more appropriate water demand relationships can be developed? A general way is to make a simultaneous specification and estimation of both demand and supply functions. Specification of a supply function, however, may be irrelevant or unnecessary for some reasons. First, a short-run supply curve of each municipality may be horizontal within a technically feasible supply range. That is, water is supplied at a fixed rate, usually irrespective of quantity used, once the price is set based in some way on supply costs. The technically feasible supply range, of course, depends also on climatic and other conditions of a particular year and time of year. Secondly, specification of a supply function does not help identifying appropriate water demand relationships, if municipalities have different demand curves as discussed before (see Figure 8(b)). Thirdly, institutional factors may be more important than physical factors *per se* (e.g. distance to a major water source) in determining observed prices and quantity of water used. For instance, municipalities' policies on water use restrictions may be quite different to each other. In some municipalities, use restrictios (seasonal or incidental) are imposed, whereas some municipality may want to sell more water for financial reasons, if population is decreasing and there is excess capacity in existing supply facilities of the municipality. Such institutional factors are difficult to reflect in supply function specification.

Probably more important is to analyse carefully the difference that appears to exist among municipalities on their demand side. A question to be asked is: Can any other variables explaining different demand behaviors of municipalities be identified or do we have to live with the fact that the demand curves are inherently different among municipalities? If the latter is true, it is necessary to treat each municipality separately or group them into classes according to their water use behavior and to study each class separately. Use of cross-section data then is further limited by number of samples that will be reduced as a result. It may be argued that cross-section studies would be more justified when a larger data base (i.e. more observations) is used for the purpose of longer-term forecasting, since the data base would probably include a wider range of municipalities representing different phases of development. A dilemma in this logic, especially when used in the context of regional planning, is that as the data base is extended to a larger area, cross-section analysis may become less relevant to a particular region of concern. Other variables may have to be included to take account of differences in geographic location, climate, social and institutional factors among municipalities.

The general inadequacy of the data base and the inappropriateness of a model developed on the basis of the cross-section data for forecasting purposes are also indicated by some water use equations identified in this study. The following water use equations were found most effective in explaining total per capita municipal water use in the year 1975.

$$y = 427.68 - \underset{(1.44^{+})}{2.334}\, x_2 - \underset{(1.26)}{63.51}\, x_6 + \underset{(3.30^{*})}{2.524}\, x_7 - \underset{(1.32)}{0.243}\, x_{12},$$

$$R^2 = 0.722,$$

$$R^* = 0.820, \quad F = 10.40.$$

$$y = 294.60 - \underset{(1.35^{+})}{2.381}\, x_2 + \underset{(5.21^{**})}{2.556}\, x_7 - \underset{(0.77)}{0.672}\, x_8 - \underset{(0.89)}{0.162}\, x_{12}$$

$$R^2 = 0.705$$

$$R^* = 0.808, \quad F = 9.58.$$

Rather than testing the models fitted to the 1975 data against the data for 1970 (another year for which extensive data are available), the regression was run using the 1970 data on the same sets of explanatory variables. The following results are obtained for the year 1970.

$$y = 219.21 - \underset{(0.87)}{2.813}\, x_2 + \underset{(0.21)}{31.72}\, x_6 + \underset{(2.60^{*})}{1.909}\, x_7 - \underset{(0.79)}{0.198}\, x_{12},$$

$$R^2 = 0.432,$$

$$R^* = 0.505, \quad F = 2.67,$$

$$y = 322.76 - \underset{(0.94)}{3.242}\, x_2 + \underset{(2.62^{*})}{1.922}\, x_7 - \underset{(0.28)}{0.498}\, x_8 - \underset{(0.57)}{0.157}\, x_{12},$$

$$R^2 = 0.434$$

$$R^* = 0.566, \quad F = 2.68 \quad .$$

The values of regression coefficients and the intercept are significantly different for the years 1970 and 1975 in the model with variable x_6 (number of person per household). Especially, the regression coefficient of variable x_6 has different sign in these equations, thus invalidating both equations as forecasting models. The model with variable x_8 is better with this respect, although the value of the multiple correlation coefficient R (or R*) is not very high for the year 1970. Other combinations of the explanatory variables were also tested, but no better result (i.e. no improvement in the R* value) was obtained.

VIII. SUMMARY AND RECOMMENDATIONS FOR FURTHER RESEARCH

Summary

The primary purpose of the present study was to show how water resources planners might analyze the municipal water demand. A statistical approach has been employed. The paper has emphasized the procedure of modelling as well as the particular results for the study region, Malmöhus County, Sweden.

An attempt to develop water demand relationships was unsuccessful. While the price variables used in the study entered the regression with positive coefficients, the validity of the models as the water demand relationships is very questionable as discussed in Section VII.

Some water use equations were identified based on the data for the year 1975 by using the total gross and the total per capita municipal water use as dependent variables. The total per capita use was fitted better with respect to residuals, which are important criteria for goodness-of-fit of a regression model.

Factor analysis was used to analyze the data base to obtain a better understanding of the demand generating factors involved in the data base. It helped to identify the appropriate sets of explanatory variables for each dependent variable.

The industrial activities were found to be most important in explaining the total per capita municipal water use. The number of people employed in water-intensive sectors connected to the municipal supply system was used as a proxy to represent this factor, since sector-wise data on industrial production were not available.

It is not entirely clear from this study alone if the number of weekend houses represents the effects of recreational activities on the total municipal water use. It may simply represent a general welfare level, since this variable and the property value have high loadings in the same factor.

No single housing variable was found very significant in explaining the total municipal water use. This is partly because the household component of the total municipal water use does not vary much among municipalities, and it is given rather as the intercept of the regression models.

Problems associated with using models developed on the basis of cross-section data for forecasting purposes were discussed in Section VII. The poor performance of the models developed based on the 1975 data on the test against the data for the year 1970 also indicates general inadequacy of the data base and/or regression models themselves.

Recommendations

Two major possibilities are seen for the extension of the present study in using the same approach and the same type of data (derived mostly from official statistics). The first one is to consider each component of the total municipal water demand separately. The second is to treat each municipality separately or to group municipalities into a few classes according to their characters and to study each class separately.

The "industrial" component, represented by the number of employees in water using sectors, was found to be most important in explaining the total per capita municipal water use. An obvious step is then to analyze the industrial water use separately. No sector-wise data, however, are available at this moment on industrial production and water use.

The first step in modelling the industrial component would be to carefully analyze the relationships among production, employment, and water use for different sectors and also for the industry as a whole. It seems to be feasible only for higher levels of aggregation than the municipality level. This analysis would indicate whether the employment figures serve as good proxy variables for industrial activities, and, if they do, how they can be used in regression.

It may be difficult to model the household component by statistical approach, due to the low variation in reported use figures among municipalities. In this case, microscopic analyses of demand generating factors may be more appropriate. That is, the amount of water used for different purposes of household operation should be analyzed for different types of housing based on the observed water use behavior. The Swedish Water and Sewage Works Association has made a review of these factors (VAV, 1975). It may be the case that the recommendations put forward on the basis of this review can serve as a sufficient proxy for household water demand. However, a careful study of the water use in weekend houses is called for.

The second major possibility listed above has an obvious problem in that the number of samples readily available is too small. Extension of the data base to include more municipalities, for example, the 277 municipalities of the whole country, has its problems as discussed in Section VII. A way to get around this difficulty is to introduce a dummy variable associated with each municipality which will take account of differences among municipalities in regression models without reducing the sample size. Another way is to use time-series data. Since availability of the time-series data is limited as discussed in Section III, pooling of cross-section and time-series data may be necessary. This poses additional problems which require further sophistication of model parameter identification techniques (Kmenta, 1971).

Another possible approach is to resort to micro data; that is, to study the basic information given in the billing records of a few municipalities of different types. This meter-based information may then be aggregated in a suitable manner, for example, in the following groups: one-family houses, multi-family houses, and industry. In this way, also the price discrimination now practiced in some municipalities may be taken into account.

REFERENCES

Andersson, A.E., P. Hjorth, L. deMaré and A.L. Thelander (1979), Water Resources Planning and Management in Advanced Economies: The Case Study of Western Skåne, Sweden - A Background Report, International Institute for Applied Systems Analysis, Working Paper, WP-79-23, Laxenburg, Austria.

Bennett, S., and D. Bowers (1976). An Introduction to Multivariate Techniques for Social and Behavioural Sciences, John Wiley & Sons, New York.

Darr, P., S.L. Feldman, and C.S. Kamen (1975), Socioeconomic Factors Affecting Domestic Water Demand in Israel, Water Resources Research, Vol. 11, No.6, pp. 805-809.

FOB 1970, 1975, Folk-och bostadsräkningen 1970, 1975 (Population and Housing Account 1970 and 1975), Central Bureau of Statistics, Stockholm.

Gibbs, K.C., (1978), Price Variable in Residential Water Demand Models, Water Resources Research, Vol. 14, No.1, pp. 15-18.

Green, P.E., (1978). Analyzing Multivariate Data, The Dryden Press, Hinsdale, Illinois.

Grunewald, O.C., C.T. Haan, D.L. Debertin and D.I. Carey (1976), Rural Residential Water Demand: An Econometric and Simulation Analysis, Water Resources Bulletin, Vol. 12, No. 5, pp. 951-961.

Gum, R.L., and W.E. Martin (1977), Structure of Demand for Outdoor Recreation, Land Economics, Vol. 53, No. 1, pp.43-55.

Johnston, J., (1972), Econometric Methods, McGraw-Hill Book Company, 2nd edition, New York.

Kim, J.R., and R.H. McCuen (1979), Factors for Predicting Commercial Water Use, Water Resources Bulletin, Vol. 15, No. 4, pp. 1073-1080.

Kindler, J. , D. Maidment and I. Gouevsky (eds.)(1980). Modelling of Water Demand, IIASA/Wiley (forthcoming).

Kmenta, J., (1971). Elements of Econometrics, The Macmillan Co.

MCB (1977), Fritidshusinventering i Malmöhus län 1976 (Weekend house inventory in Malmöhus County 1976), Malmöhus County Board, Malmö.

McCuen, R.H., W.J. Rawls and B.L. Waley (1979), Comparative Evaluation of Statistical Methods for Water Supply Forecasting, Water Resources Bulletin, Vol. 15, No. 4 pp. 935-947.

Morgan, W.D., (1973), Residential Water Demand: The Case from Micro Data, Water Resources Research, Vol. 9, No. 4, pp. 1065-1067.

SCB 67-77, Årsbok för Sveriges kommuner (Statistical Yearbook on Swedish Municipalities), Central Bureau of Statistics, Stockholm.

Sewell, W.R.D., and S.B. McClellan (1979), Refinement of Techniques for Water Demand Forecasting: A Review of Some Canadian Contributions, International Institute for Applied Systems Analysis, Working Paper WP-79-3, Laxenburg, Austria.

Sonnen, M.B., and D.E. Evenson (1979), Demand Projections Considering Conservation, Water Resources Bulletin, Vol. 15, No. 2, pp. 447-460.

SOU 1965:8, Skånes och Hallands Vattenförsörjning (Water Supply of Skåne and Halland), Ministry of Communication, Stockholm.

Turnovsky, S.J., (1969), The Demand for Water: Some Empirical Evidence on Consumers' Response to a Commodity Uncertain in Supply, Water Resources Research, Vol. 5, No. 2, pp. 350-361.

VAV (1975), Vattenprognos 1975-2000 (Water Use Forecast for 1975-2000), Swedish Water and Sewage Works Association, VAVP30, Stockholm.

VAV, S67-77, VA-verk 1967-1977 (Statistics on municipal water and seawage works 1967-1977), Swedish Water and Sewage Works Association, Stockholm.

VAV, Tx69-77, VA-taxor 1969-1977 (Municipal Water and Sewage Charges 1969-1977), Swedish Water and Sewage Works Association, Stockholm.