TIPIFICAÇÃO DE CAUDAIS DE REDE EM SISTEMAS DE DISTRIBUIÇÃO DE ÁGUA

Dália Loureiro [1], Sérgio Teixeira Coelho [2], Paulo Machado [3], Alexandre Santos [4], Helena Alegre [5], Dídia Covas [6]

[1] *Bolseira de Doutoramento, Laboratório Nacional de Engenharia Civil, dloureiro@lnec.pt*
[2] *Investigador Principal, Laboratório Nacional de Engenharia Civil, stcoelho@lnec.pt*
[3] *Investigador auxiliar, Laboratório Nacional de Engenharia Civil, dloureiro@lnec.pt*
[4] *Consultor, WADI- Consultoria Unipessoal, Lda,* alwadi@sapo.pt
[5] *Investigadora Principal, Laboratório Nacional de Engenharia Civil, halegre@lnec.pt*
[6] *Professora Auxiliar, Instituto Superior Técnico, dídia.covas@civil.ist.utl.pt*

Resumo

Compreender os factores que influenciam os padrões de consumo diários e semanais em sistemas de distribuição de água é fundamental para actividades a longo prazo, como seja o planeamento, projecto, expansão e reabilitação dos sistemas, e para actividades a curto prazo, como sejam actividades de rotina no âmbito da exploração dos sistemas ou da gestão de emergências.

Este artigo apresenta um programa nacional para a caracterização de consumos domésticos, através do uso de um protocolo específico para a recolha e análise de dados de consumo e de uma base que reúne um conjunto de atributos, potencialmente relacionados com o consumo relativos à rede (e.g., pressão de serviço, estado e nível de manutenção da rede, utilização de reservatórios domiciliários), ao sistema de facturação e de gestão de clientes e ao perfil sócio-demográfico dos clientes abastecidos (e.g., idade, nível sócio-económico, mobilidade, idade e tipo de edifício, etc.).

O programa está actualmente em curso tendo por base dados relativos a um conjunto de cerca de 20 sectores de rede, operados pelas 9 entidades gestoras participantes no projecto INSSAA, e que compreende redes com características variáveis, em termos de dimensão, do comportamento hidráulico e das características de consumo. A dimensão dos sectores varia entre 2000 e 12000 ramais ligados e a maioria dispõe de um programa de monitorização de caudais desde 2005 (em alguns casos desde 2004). O consumo total entrado nos sectores de rede é monitorizado em contínuo, dispondo-se, na maioria dos casos, de registos durante as épocas de Verão e de Inverno.

Obtiveram-se padrões de consumo normalizados e estatísticas de consumo para cada dia da semana, assim como para cada estação, no caso de se tratar de área com efeito de sazonalidade. A informação sócio-demográfica sobre a população abastecida, em cada sector, foi obtida a partir do último recenseamento geral da população e da habitação (Census, 2001) ao nível da subsecção estatística.

Após esta fase inicial, o programa será disseminado e tornado acessível a todas as entidades gestoras que desejem participar, num princípio de voluntariado e mantendo toda a informação anónima. O acesso a esta base de dados será feito via web, através de um webiste dedicado, com mecanismos de consulta e de análise, assim como de um mecanismo que permita adicionar novos casos de acordo com o formato especificado pelo protocolo. A publicação dos resultados desta base de dados fornecerá aos projectistas, consultores e gestores informação detalhada e fiável sobre consumos de água até agora inexistente. Permitirá também estimar o consumo, a partir de atributos conhecidos, em sectores onde não existe ainda um histórico de caudais.

Este artigo apresenta a metodologia preconizada no âmbito deste protocolo, assim com o conjunto de variáveis potencialmente relacionadas com o consumo, construídas com base na informação recolhida, e ilustra também os resultados preliminares obtidos.

Nota sobre o artigo: Este artigo tem por base uma comunicação apresentada na conferência sobre “8th Annual International Symposium on Water Distribution Systems Analysis”, Cincinnati, Ohio, E.U.A., 27-30 de Agosto, com o tema “Profiling residential water consumption”.

Keywords

Water consumption; demand patterns; demand analysis; explanatory factors

1 INTRODUCTION

Detailed knowledge of water demand statistics and demand patterns is required by many situations involved in the planning, design and operation of water supply and distribution systems. Often there are insufficient data available for the particular area and it is necessary to estimate based on data from other similar networks, or on general, often artificial statistics such as per capita values, peak factors and literature demand patterns. In Portugal, as in many other countries, there is a lack of systematic studies and data that may enable designers and network managers to infer useful information about behaviors and demands based on the geographical and socio-demographic contexts and network type.

This paper describes the development of a Portuguese nationwide program for characterizing residential water consumption, through the use of a specific consumption data analysis protocol and the build-up of a database of demand patterns; and for associating it with a range of easily obtained, potential explanatory factors, such as technical features of the network, billing statistics and social-demographic parameters. The objective of this program is two-fold: to develop and disseminate a systematic and sound procedure to characterize urban water consumption; and to create a data base that allows for extrapolation in cases for which no information about consumption is available. The protocol defines, among others, the requirements for the metering districts considered (topology, homogeneity of residential demand types, monitoring points, metering hardware, etc), for the flow data to be recorded (time step, units, etc), and for the context information.

Besides the database in itself and the search for correlation parameters, the program aims to explore query mechanisms that facilitate the profiling of projected consumption requirements and the search for similar cases.

This program draws on experience collected through previous studies of consumption behavior undertaken in the cities of Lisbon and Almada, in Portugal (Alegre, 1992; Alegre and Coelho, 1993), and was made possible by a number of recent developments, including the increased availability of district metering data across the country; a growing awareness of the need for water accountability; emerging technologies such as household / small district consumption telemetry; and significant improvements in the data and digital tools available for spatial analysis of census data in Portugal (Census, 2001).

The program is currently being test-run on a set of 20 metering districts from 9 water utilities, drawing from a pool taking part in a collective network modeling initiative (INSSAA, 2006). They comprise a variety of network types and demand characteristics across representative urban areas of Portugal. The test metering districts range in size from 2,000 to 12,000 connected properties, and are being monitored since 2005 (in some cases, 2004).

After the initial run, the program is to be disseminated and made accessible to all water utilities willing to participate, on a voluntary basis and with data anonymity. This will be done mostly through a dedicated website containing the consumption database, together with query and analysis tools, as well as a supervised mechanism to submit new cases in standard format. It is believed that the publication of the result database and its continued growth will provide the water industry designers, consultants and managers with much more reliable and updated data on water consumption than previously available. It will also represent a means to estimate demand through correlation and profiling, in cases when there are no records to work from; and to assess the impact of certain types of network alterations or socio-demographic evolution on the behavior of demands for a particular network.

The paper discusses the demand analysis performed, as well as the development of the set of potential explanatory factors – technical features of the network and social-demographic variables – and illustrates with results obtained so far.

2 METHODOLOGY

2.1 General approach

The methodology followed in the study includes the following steps:

i) Continuous recording of flow consumed at residential network districts;
ii) Collection of static information to support the analysis (billing information; network data; social-demographic information);
iii) Statistical modeling of consumption at metering district level in order to obtain standardized consumption patterns and relevant volume statistics;
iv) Selection of a group of technical and social-demographic variables that potentially influence water consumption;
v) Analysis of detectable relations between these variables and water consumption levels and daily/ weekly/ seasonal patterns.

The different types of data to be collected are detailed in section 2.3.

Total consumption is continuously monitored during the main seasonal scenarios (typically winter and summer) in the areas selected. Standardized daily patterns and consumption statistics are produced for each day of the week (as well as for each season, in case there is seasonality in the area). Social-demographic information is based on statistics obtained from the 2001 National Census for the basic statistical units, which are around 300 dwellings in size.

The variables that are considered for potential correspondence include:
- Technical features of the network, such as mean service pressure and pressure variation, network state and maintenance level, possible use of household tanks;
- Billing and customer statistics; and
- A range of social-demographic variables, such as consumer age, social-economic grouping and mobility, age and type of dwelling, and economic activities.

In essence, the program aims at collecting information and data so as to characterize each metering district in a standardized way. The main elements that make up this profile of a metering district are of 4 types:
- Dimensionless daily demand patterns, on a given standard time grid;
- Consumption statistics (e.g., max, min and mean daily flow);
- Technical explanatory variables
- Socio-demographic explanatory variables

Section 3 describes the main aspects to be taken into account as regards the computation of daily patterns and consumption statistics, and section 4 looks into the selection of technical and socio-demographic driver variables that may potentially be related to water consumption.

2.2 Metering district selection

The metering districts to be considered must fulfill a number of requirements, whose details are described in Table 1.

Table 1. Requirements for metering district selection.

Requirement	Description
Flow monitoring conditions	– Total inflow must be metered through reliable and adequately sized and fitted flowmeter(s). – Minimum duration of monitoring campaign: 2 months.
System operation	– Boundaries of the network district must be clearly defined and watertight (either permanently or for the duration of the campaigns). – Operating configurations must be recorded and kept constant between different metering campaigns (such as, e.g., for different seasons) to ensure data consistency.
Consumption	– The metering district must be composed predominantly of residential users for each consumption scenario. – Any large consumers (i.e., those whose consumption is likely to have a significant effect on total inflow and its daily behavioral pattern) must be individually metered.
Type of dwelling	– Percentage of private household dwellings greater than 50% of the total in the district.
Type of building	– Percentage of residential buildings greater than 50% of the total in the district.

Although not a vital requirement, the availability of network analysis models for the districts considered has definite advantages, given the improved understanding of the hydraulics (especially pressures) and the possibility of scenario or hypotheses testing. Conversely, this type of consumption studies greatly benefits the modeling effort, through better specification of demands and demand patterns.

2.3 Data collection

2.3.1 Geographical data

Geographical data comprise data about the water distribution network itself, and the cartographic and socio-economic geo-datasets of the territory it is located on.

Geographical data about the network aims to characterize system components and system hydraulic behavior, in this occurrence mostly as a means to estimating the real losses component of the total water use. The collection of this type of data requires an updated GIS system and, as mentioned above, benefits from the availability of a calibrated network model for the estimation of daily pressures, a key element in the estimation of losses. Mains and household connection failure indicators were based on IWA indicators (Alegre *et al.*, 2006) and provide important information for the characterization of network condition.

The cartographic and socio-economic geo-datasets of a metering district are mostly used for generating additional information about the type of consumers, water use and occupation, helping in the identification of homogeneous areas derived from population and their specific socio-economic profiles. This is possible using the statistics micro-area sub-section, or SSE (*"Subsecção Estatística"*) from the 2001 Census, as made available by the National Statistics Institute (INE). A SSE is the smallest territorial unit for statistical use available in Portugal and corresponds to the smallest homogeneous building and living area existing inside a statistics section (a continuous area of one parish, with about 300 dwellings). Spatially, SSE are polygons ranging from less then 1 ha, in urban areas, to 10-15 ha, in low density populated areas. In general terms, a SSE corresponds to a block in urban areas (Census, 2000), and to a place or part of a place in rural areas. To give a measure of scale, the combined total of 42,000 SSE in the metropolitan areas of Lisbon and Porto averaged in the 2001 census 18 buildings (median = 12), 46 dwellings (median = 22) and 101 residents (median = 51) per SSE. Each of these SSE polygons have demographic and socio-economic data associated, covering more than 50 different attributes.

Cartographic data include data from municipal databases, namely aerial images, road networks, address numbers, buildings, among others, allowing the identification of metering district boundaries (with pipe network and household connections). It is mostly used for generating additional information about the type of consumers, water use and occupation; to aid in the definition of metering district boundaries, integrating network data (such as mains, household connections, etc), demographic and SSE components using GIS tools, as described on section 2.3.4 (see Figure 1); and to cross-validate data from different dates.

Table 2. Geographical data to be collected for each metering district.

Data	Data requirements	Data source
Network data		
Pipe length, material, diameter and age	– Total pipe length – Georeferenced pipes – Pipe failure indicator (nº/100 km/year)	GIS/ Network manager
Household connections (total number)	– Total number of household connections – Georeferenced household connections – Household failure indicator (nº/1000 household connections/year)	GIS/ billing system
Network model	– Calibrated network model	Network model
Data about territory metering district		
Household tanks	– Number of users supplied by household tanks; tank volume, emptying/filling schedule	Network managers
Buildings and addresses location point	– Georeferenced data	GIS
Aerial images	– Georeferenced data	GIS
Street lines	– Georeferenced data	GIS

2.3.2 Billing and customer data

Billed consumption is metered or estimated monthly, bi-monthly or half-yearly, and provides the following relevant information:

- Total billed consumption and total number of customers, per category of consumption, for the consumption scenarios considered;
- Total non-revenue water in the metering district (by comparing with the network input volume) and;
- Relative weight of categories of consumption for each scenario.

Billing and customer data provides information about the importance of residential consumption relative to other categories of consumption (e.g., commerce and industry, collective, municipal) in the total billed consumption of a metering district. This complements the social-demographic analysis, as it allows for the quantification, in terms of billed consumption, of the degree of homogeneity of residential consumption in the district. Social demographic data focus only on resident population, which corresponds to the residential consumption category (though not totally: it also accounts for temporary residence and some family-oriented services). Billed consumption can provide also important information about seasonal effects (e.g., municipal uses such as watering of public green spaces may represent a higher fraction of the total in summer than in winter, while domestic consumption may actually decrease in summer due to the vacation period) – it is important to note the monthly variation of consumption and not just the annual totals. The spatial distribution of billed consumption (only possible if georeferenced at household connection level or similar) is valuable for the analysis of spatial heterogeneity of consumption. Billing data also provide a global appreciation of the significance of large consumers, in terms of average consumption and annual variation.

Non-revenue water is composed of unbilled authorized consumption, apparent losses and real losses. Unbilled water comprises all the authorized uses of water related to public purpose (e.g. fire fighting) or to the undertaking's own consumption that are nor billed. Apparent losses can be grouped into unauthorized consumption and metering inaccuracies (IWA, 2000). Unauthorized consumption may be due to different factors (e.g., meters that have been tampered with; illegal use of hydrants; illegal connections) and is estimated based mostly on network manager insight. Meter inaccuracies occur mainly when low flow conditions take place, or with meters in poor condition. Real losses may be assessed as the difference between non-revenue water and the sum of unbilled authorized consumption and apparent losses, or estimated based on minimum night flow analysis (Garcia *et al.*, 2003; WRC, 1994).

Compared to traditional meter reading, the emerging household telemetry metering technology facilitates the data processing and improves the accuracy of water balance estimation.

Table 3. Billing and customer data to be collect per metering district.

Data	Data requirements	Data source
Categories of consumption	– Categories of consumption with detailed uses per category – Number of customers per category of consumption	Billing and customer system
Uses per category of consumption	– Uses per category of consumption (e.g., commerce and industry can include shops; supermarkets; bars, restaurants, warehouses; factories)	Billing and customer system
Billed consumption per category	– Total annual billed consumption – Annual variation of consumption – Georeferenced billed consumption (by household connection)	GIS/ Billing and customer system
Large consumers	– Total annual billed consumption – Annual variation of consumption	Billing and customer system
Unbilled authorized consumption	– Unbilled unmetered consumption from authorized residential and businesses uses – Unbilled watering of gardens consumption – Unbilled street cleaning consumption – Unbilled fire fighting – Undertaking self consumption	Network manager
Unauthorized consumption	– Estimated value	Network manager
Metering inaccuracies (sub-metering)	– Estimated value	Network manager
Real losses	– Leakage on transmission and/or distribution mains, leakage and overflows at transmission and/or distribution storage tanks, and leakage on service connections up to the measurement point	Network manager

2.3.3 Network flow data

The inflows into metering districts should be monitored through properly calibrated flowmeters with adequate flow ranges and error curves. Pressure metering is carried out near the supply point and at a least favorable node, often at the opposite end of the network or at the higher elevation consumption locations. The objective is to assess service pressures as well as total head loss and. Pressure records are an important aid in validating flow data at the analysis stage.

Before starting data recording, a set of procedures are required, such as:
i) Checking boundary valves in order to confirm the correct isolation of survey areas;
ii) Informing maintenance staff about which valves are to be kept temporarily closed;
iii) Assuring that either there are no maintenance works that may affect the results, or that any works are adequately documented for future reference.

A preliminary campaign should be carried out to check the whole monitoring system. During the monitoring periods, a supervisor will make sure flow and pressure data recording starts and stops according to scheduling. Immediately after campaign end, data should be carefully observed in order to detect and correct in due time any deficiency that may occur.

The flow records generated are the central piece of data in the process. Their processing and analysis is described in section 3.

2.3.4 Social-demographic data

It is important to establish a model of social characterization while trying to understand the key driving factors that influence daily and weekly water consumption patterns in distribution systems. This is because water consumption varies according to the social profile and the demographic characteristics of the residents. Their behavior in situ and attitude toward water consumption finds a high level of explanation in this type of interdisciplinary applied research. Additionally, it may be easier to get information on the social demographic characteristics than on the water consumption, especially in expansion areas. This information can be used to estimate water demand profiles, if the referred model is available.

Previous research has shown that consumption is influenced by the social-demographic characteristics of consumers (Alegre and Coelho, 1993; Alegre *et al.*, 2005). For example, Burnell (2003) and Alegre and Coelho (1993) showed that areas with a high percentage of retired citizens present an almost constant daytime water use pattern, due to a well settled routine and long periods at home. Domestic habits are based on a social regularity condition, which supports the notion that social-demographic characteristics may relate well with water network consumption.

Like so many other technical domains (marketing is a good example), the social stratification of water consumers is currently placed at the same level used to promote the social knowledge of a population, such as variables like gender, nationality, region of birth, age, revenue or profession, among others. However, social stratification is slightly more complicated to determine because it is necessary to use sociological indicators in order to infer a certain position in a given social structure. This procedure should be a quantitative approach of a theoretical social stratification model, but the data requested is obtained on behalf of census national operations (which is by nature, static and not frequently updated).

Social-demographic variables are in this case provided by GIS data obtained from the last national census on population and dwelling (Census, 2001), made available at the SSE level by the National Statistics

Institute (INE). Social-demographic data collection aims at producing indicators potentially associated to daily or seasonal consumption behavior, from a set of census variables. The overall social-demographic characteristics of a metering district will be obtained based on the dominant social-demographic characteristics of each SSE included in the metering district.

Figure 1. Metering district boundary using SSE and household connections.

The set of socio-demographic variables is divided into the following groups (broad categories of information) directly resulting from the last Census (2001):

– Building
– Dwelling
– Private household
– Household status
– Individuals

Since the variables included in “Private household”, “Household status” and “Individuals” consider only the resident population, the contribution of social-demographic indicators for the explanation of consumption patterns increases as the weight of domestic consumption increases in the total consumption of a metering district.

In order to obtain social-demographic indicators for each metering district, an application has been developed in ArcGIS ® that crosses network data (i.e., mains, connections, network district boundaries) with social-demographic data, aggregated at SSE level, obtained from INE in GIS-ready format. The SSE that are intersected by the metering district boundaries are weighed by surface area, assuming homogeneity applies. The procedure is as follows:

i) Data collection;

ii) SSE identification (by matching geographic data with social-demographic data, see Figure 1);
iii) Identification of the main variables that should be taken into account for the SSE's social-demographic characterization;
iv) Assessment of social-demographic variables for each SSE;
v) Establishment of the metering district profile trough the analysis of SSE.

Household connection georeferencing is very important for this procedure, because it allows for:

- Locating the population supplied by the metering district, establishing the SSE that make up the metering district (see Figure 1) and calculating the number of users per household connection and their social-demographic characteristics – age, economic activity, average family size, permanent residence or weekend/holiday dwelling, among others;
- Detection of incoherent data (e.g., a new block of buildings built after the national census in a SSE without population) and;
- Computation of total billed consumption per sub-statistical unit (in the utilities where it is possible to connect GIS and billing system).

Some difficulties arise due to the following:

- Geographic elements, including data census, may have different timestamps, and making them compatible can be time consuming;
- National statistics at SSE level are produced only once each decade, limiting the use of this methodology for rapidly growing urban areas;
- Although a number of water utilities covering a significant population universe have georeferenced household connections in their mapping systems, many water utilities in the country still do not have this technology; other geographic elements should be used for estimating supplied population, such as address location point.

2.4 Standardized data collection protocol

The protocol for data collection and characterization of consumption at metering district level involves the definition of a data base that can conciliate the data requirements described in 2.3 with the data available at the majority of water utilities. A preliminary survey of the most important data constraints has shown that the establishment of a standardized protocol involves the following challenges:

- Use of data provided by multiple information systems (GIS, billing and customer system, flow monitoring system, national census);
- Importance of georeferenced data that can be explored together to produce useful statistics;
- Importance of a good district metering and flow data recording system.

Such a protocol is essential in order to guarantee the consistency of the information and of the analysis results. The protocol has been materialized in a data collection form and database in MS Access®.

3 NETWORK FLOW DATA: STANDARDISATION AND ANALYSIS

3.1 Daily consumption patterns

Urban water consumption has a stochastic behavior where three cyclical patterns may be identified: a daily pattern, a weekly pattern and a seasonal pattern. On top of these, multi-annual trends may occur (i.e., the time series may be non-stationary). Seasonal effects may or may not be present; in Portugal, they tend to be more noticeable in seaside areas, mostly due to tourist influx. In most urban or metropolitan areas, the effect tends to be less perceptible, and is largely linked to increased watering of green areas. In some way, the vacation period of the residents is compensated by visitors. The occurrence of seasonal effects determines whether the particular metering district is profiled for a single scenario or multiple scenarios.

Daily consumption patterns are a prime tool for studying and translating user behavior and water use variations throughout the day. Network modeling and several other engineering and planning analyses depend heavily on this key element and associated flow statistics, e.g. peak factors. The production of such daily consumption patterns is based on the statistical processing of continuous network flow records, in order to obtain some sort of average curve over the 24 hours of the day.

A daily pattern is no more than a curve uniting some representative value of consumption at each successive point throughout the 24 hours of the day (Figure 2). Different days of the week may correspond to different patterns. Each point represents a population of all the consumption values recorded at that particular time of the day, and day of the week. If long-term trends are removed, consumption behaves as a simple random variable. Statistics of random variables may then be applied, in order to test if characteristics and trends identified in the observation stage are meaningful. Coelho (1988), using consumption data from urban areas near Lisbon, Portugal, has shown that these cross distributions tended to present several peaks and be rather scattered. According to the same research, this effect seems to disappear when raw flow data is divided by its corresponding daily average, producing a dimensionless pattern. The log normal distribution provided in that case the best overall fit for the different times of the day (the populations tend to show a non-simmetrical shape for low flow times, and symmetrical for the remainder). De Marinis *et al.* (2006) have suggested the use of a number of different distribution functions (including the log-normal, Gumbel, normal and Poisson models) according to the time of the day. Given its versatility in describing non-simmetrical populations, the log-normal was retained for the computation of daily patterns in this study.

It is beneficial to complement the mean curve with confidence bands that give an idea of dispersion at each point. Removing the flow dimension by dividing by the daily average has certain advantages: seasonal effects due to population fluctuations are felt on flow values but not significantly on daily behavior of the curve. The dimensionless pattern is thus less influenced by those seasonal effects. When these become pronounced – which can be seen by testing the series of daily totals, for example – it is preferable to divide the analysis into more homogeneous periods, representative of each season.

In this program, the recommended flow monitoring frequency is of 10 to 15 minutes, recording either total or average flow over the time step. That frequency is thought to be sufficiently detailed for a correct depiction of daily variation and identification of a good set of behavior features. Shorter time steps will yield higher detail but also increase data storage requirements dramatically. One limitation of using a 10 to 15 mins time step is the difficulty in estimating real losses from night flows, as the time step is too coarse to assess the actual leakage background flow.

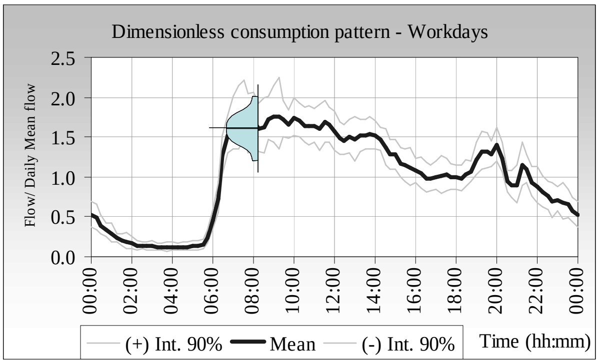

Figure 2. Standardized dimensionless daily consumption pattern.

One aspect to be tested is the difference of daily consumption patterns between workdays. Hypotheses tests may be performed using minimum night flows, average consumption and peak consumption time series. The shape and intensity of the dimensionless diagrams should also be taken into account. For most of the districts surveyed so far, no meaningful difference has been found between workdays, with the exception of leisure weekend areas in the vicinity of metropolitan areas. Saturdays and Sundays tend to be different from one another and from workdays.

3.2 Network flow statistics

Based on daily consumption patterns, the following procedure to obtain network flow statistics for each metering district is proposed:

i) Estimation of different components of water balance;
ii) Estimation of daily patterns due to domestic consumption;
iii) Calculation of network flow statistics that characterize daily consumption patterns;
iv) Classification of daily consumption patterns.

Regarding the different components of water balance (Alegre *et al.*, 2005; IWA, 2000), total network flow can be decomposed into billed authorized consumption, unbilled authorized consumption, apparent losses and real losses. Social-demographic analyses focus on the characterization of household users whose consumption (metered or estimated) can be included in the billed and unbilled authorized consumption.

Therefore, it is important to estimate the different components of water balance, according to information collected in 2.3.2, for each consumption scenario, in order to isolate domestic consumption:

- *Non-revenue water*;
- *Unbilled authorized consumption;*
- *Apparent losses*;
- *Real losses*: estimated based on minimum night flows analysis (Garcia *et al.*, 2003; WRC, 1994) and assumed constant during the day;

The component of real losses is assumed constant during the day whenever the daily pressure variation is insignificant in the metering districts analyzed (see 4.1). Network flowmeter inaccuracies are assumed negligible, since it is assumed they were correctly sized and installed, which is a basic requirement for metering district selection (see 2.2).

After estimation of domestic daily consumption patterns, it is possible to obtain network flow statistics according to Table 4, which characterize daily, weekly and seasonal consumption behavior.

Table 4. Network consumption statistics.

Daily statistics	Daily average consumption Daily minimum consumption Daily peak consumption Daily behavior - analysis per periods of day of average consumption, consumption variation
Weekly statistics	Weekly peak consumption Identification of statistics that allow for the detection of differences in daily behavior between workdays, Saturdays and Sundays
Seasonal statistics	Seasonal peak consumption Identification of statistics that allow for detecting seasonal differences in daily behavior.

After the calculation of network consumption statistics, the daily consumption patterns are classified.

4 SEARCH FOR DRIVING FACTORS

4.1 Network statistics

Network statistics aim at providing information for the estimation of the real losses component and detect potential influence of pressure on consumption:

- Network status
- Daily average service pressure across the metering district
- Daily pressure variation at the least favorable consumption point.
- Excess pressure ratio (ratio between daily average pressure and minimum pressure required)

4.2 Billing and customer-related statistics

Billing and customer-related statistics allow for the characterization of different categories of consumption. Although the predominance of residential users constitutes a requirement for metering district selection, it is important to analyze the importance of other categories per scenario of consumption.

In general, the customer categories and the types of consumers included in each category vary between water utilities. Categorization of customers is not oriented by water uses, but as a function of tariffs (e.g., primary schools can be found in the Municipality category and secondary schools can be found in the State category). Therefore, in order to compare categories of consumption between metering districts, the number of customer categories was reduced and standardized. The following customer categories are adopted in this program:

- Residential customers (may include non-residential customers that have similar uses to the residential category);
- Commercial and industrial customers (e.g., shops, supermarkets, restaurants, warehouses, industries, etc.);
- Collective customers (e.g., schools, public buildings, hospitals, sports facilities);
- Municipal customers (e.g., watering of public gardens, fountains, kindergartens).

The following statistics have been established for each consumption scenario:

- Weight of each category of consumption on total billed consumption
- Weight of large consumers per category of consumption and on total billed consumption
- Per capita consumption (or equivalent per capita consumption) for each category

4.3 Social-demographic statistics

From the set of social-demographic variables provided by the National census, the following variables have been considered as starting point variables for the analysis, based on the experience collected so far and on previous studies (Alegre, 1992; Alegre and Coelho, 1993):

i) Total number of residents;
ii) Total number of dwellings (includes non authorized dwellings);
iii) Total number of classic dwellings (excludes non authorized dwellings);
iv) Total number of classic dwellings with water supply;

v) Total number of exclusively residential classic dwellings;
vi) Total number of classic dwellings used as 2nd residence/holiday residence;
vii) Age group – residents over 65 years old;
viii) Number of residents working in Trade, Tourism and Services (the Tertiary Sector of Activity);
ix) Number of private households;
x) Number of private households without unemployed members;
xi) Number of residents with a university degree;
xii) Number of residents that concluded up to the 4th grade.

This initial list of variables may be adjusted as the social-demographic analysis develops for different metering districts, in the scope of the on-going tests. Census variables will be used to produce indicators for the classification of each sector, according to the most probable points of view that can influence consumption habits: e.g., social-economic, building, dwelling, mobility, seasonality.

Considering the social-economic point of view, the following indicators were derived from items vi) to xii):

- *Ageing Index* (AI) – percentage aged 65 plus, considering the total of residents;
- *Economic Activity Changing Index* (EACI) – percentage of residents working in Trade, Tourism and Services (tertiary activities) within the total of residents employed;
- *Vulnerability to Unemployment Index* (VUI) – percentage of private households with (at least) one of its members unemployed , within the total of private households;
- *Social Status Ratio* (SSR) – Ratio between the number of adult residents with a university degree and the number of adult residents that did not continue their studies beyond 4th grade.

These social indicators were previously tested in a sample of over 42,000 SSE, in order to determine the main grouping for each (see Table 5).

Table 5. Measures of social-economic indicators (expressed in %).

Indicators	1º quartile	2º quartile	3º quartile	4º quartile
[AI] Aging Index	**≤ 11.3**	**> 11.3 ; ≤ 17.2**	**> 17.2 ; ≤ 24.6**	**> 24.6**
	LOW	TENDENCY TO LOW	TENDENCY TO HIGH	HIGH
[EACI] Economic Activity Changing Index	**≤ 24.5**	**> 24.5 ; ≤ 38.1**	**> 38.1 ; ≤ 57.7**	**> 57.7**
	LOW	TENDENCY TO LOW	TENDENCY TO HIGH	HIGH
[VUI] Vulnerability to Unemployment Index	**0**	**> 0 ; ≤ 7.7**	**> 7.7 ; ≤ 13.3**	**> 13.3**
	LOW	TENDENCY TO LOW	TENDENCY TO HIGH	HIGH
[SSR] Social Status Ratio	**≤ 6.1**	**> 6.1 ; ≤ 13.6**	**> 13.6 ; ≤ 35.1**	**> 35.1**
	HIGH	TENDENCY TO HIGH	TENDENCY TO LOW	LOW

Other statistical tests were developed with the goal of understanding the level of correlation between these indicators with others (available for municipalities but not at the SSE level). Considering the good results obtained, proving that these four indicators are really useful substitutes of data related to social stratification, a specific scale of social stratification (SSS Index) was developed, according to the following equation:

$$SSS = (3 \cdot SSR) + (2 \cdot VUI) + 1.5 \cdot (EACI) + AI$$

Higher values of this index mean a higher social position. Each SSE and the metering district (as the unit of analysis) can be categorized according to their SSS score.

5 RESULTS AND DISCUSSION

The objective of this study is to profile water consumption from distribution networks and relate it to network, billing and customer, and social-demographic statistics, in order to allow for extrapolation for cases for which reliable consumption patterns are not available.

Concerning the latter, Figures 3 and 4 illustrate two examples of intersection of network data with social-demographic data, aggregated at SSE level, for a metering district. Figure 3 shows the spatial distribution of exclusively residential buildings. Although in most of the SSEs the buildings are mainly residential, it is important to realize the importance of the remaining SSEs on consumption which have a reduced percentage of residential buildings (in most of them equal to zero). Figure 4 illustrates the spatial distribution of individuals older than 65 years (retired individuals). Based on this type of variables it is possible to establish indicators such as the Aging Index, the indicator presented in 4.3 for social-economical characterization.

Figures 5 and 6 illustrate how consumption patterns can be associated to social-demographic habits. At this stage, the losses component has not been deducted from the patterns showed.

Figure 5 shows daily patterns for an area with an average social-economic level. Daily patterns are marked by a high variation of consumption throughout the day. Workdays are characterized by peak consumption at 7:30 am and it can be observed that the diurnal consumption (between 8:00 am and 6:00 pm) is significant. For the majority of families, both members of the couple have outside jobs; the entire family is out early and only returns in late afternoon. The period of higher consumption between 9:00 am and noon may be due to house cleaning by housekeepers; lunch time is not noticeable. Weekend consumption patterns are marked by a different profile from workdays. On Saturdays and Sundays the first peak consumption occurs at 10:00 am and the second peak consumption occurs near 12:00 pm. The confidence intervals display a large range, which can indicate more heterogeneous behaviors between consumers. This metering district is marked by reduced night flow values. Daily mean flow values reveal a regular consumption during weekdays. However, this metering district was monitored during summer and it was observed that weekend consumption decreases comparatively to workday consumption, which can be explained by the fact that a significant percentage of the families go out for the weekend.

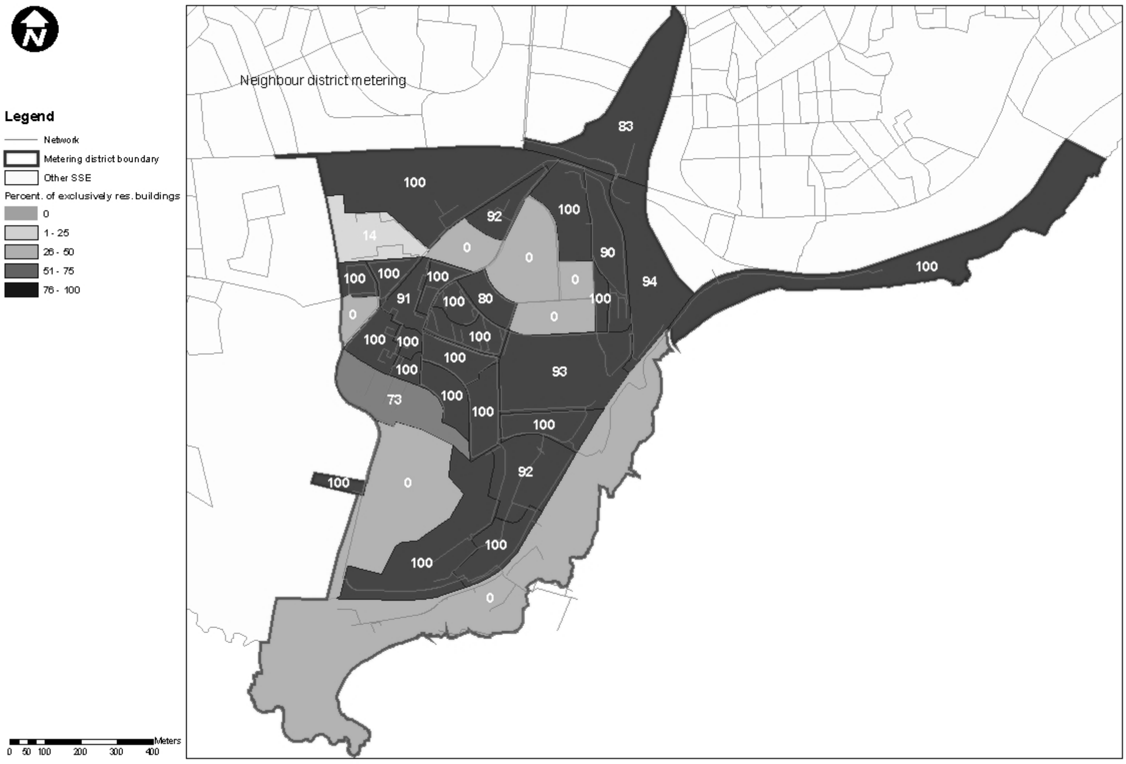

Figure 3. Spatial distribution of the '*exclusively residential buildings*' variable.

Figure 6 shows daily patterns for an area with a low social-economic level, calculated over the same period as Figure 5. The workday pattern is flatter during daytime, showing a local peak near lunch time (noon) and an absolute peak near 8:00 pm, which coincide with the main meals. This is consistent with the fact that for a significant percentage of couples in this district, the wife works at home. There is also a higher percentage of unemployed. The Saturday and Sunday consumption patterns are rather different between them and different from workdays. On Saturdays, the daily pattern registers important consumption between 12:00 pm and 6:00 pm which can be due to uses other than direct human consumption, such as cleaning houses, car washing or garden watering. Analyzing daily mean flow values, it can be seen that maximum values take place on Saturdays. On Sundays the peak consumption values take place near noon and around 8:00 pm, which is a behavior more similar to the one shown in Figure 5.

Despite the differences in the behavior of daily consumption patterns for the two situations with different social-economic levels, the per capita volumes of water consumed are overall equivalent.

Figure 4. Spatial distribution of the ‘*individuals aged over 65 years*‘ variable.

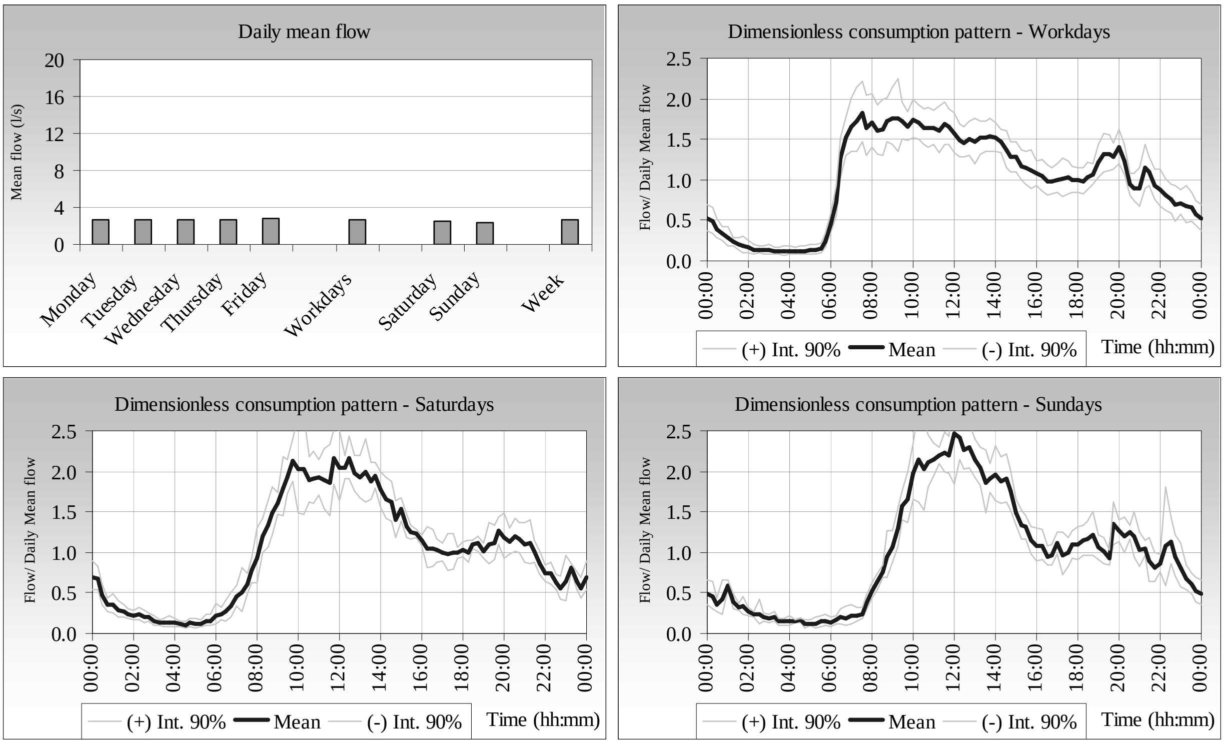

Figure 5. Daily consumption patterns for an area with average social-economic level.

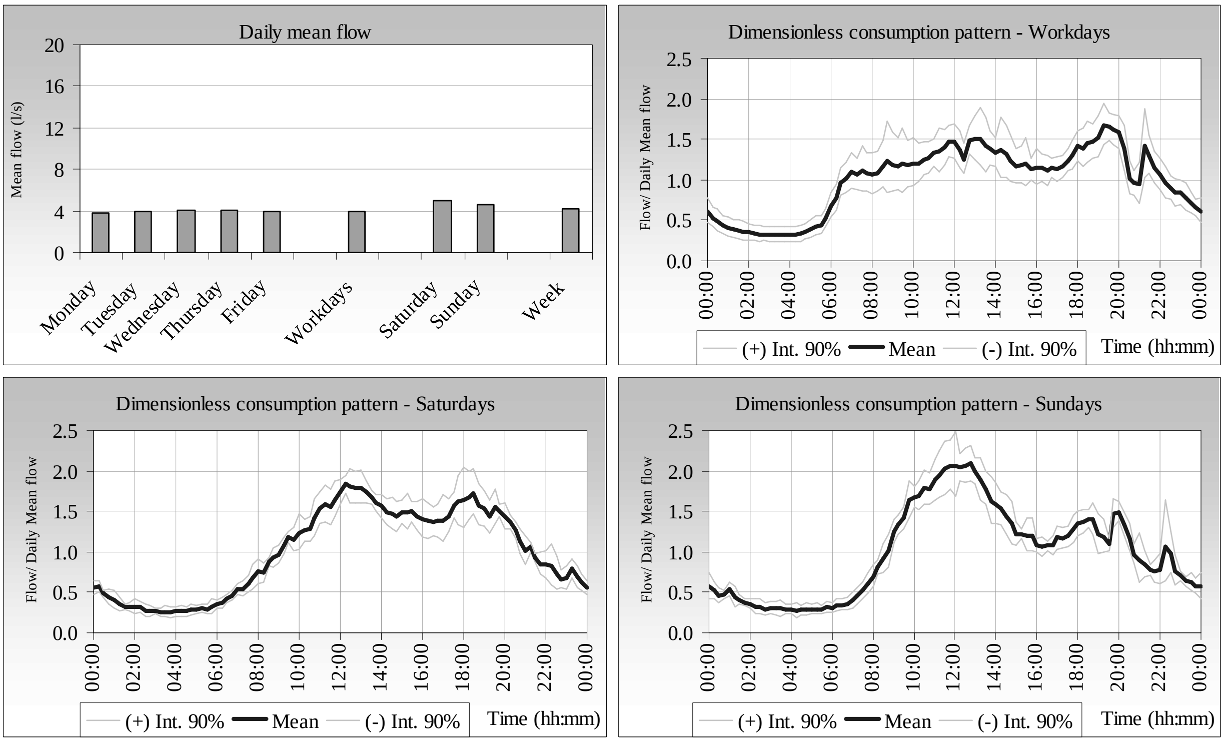

Figure 6. Daily consumption patterns for an area with low social-economic level.

6 CONCLUSIONS

The development of a program for characterizing residential water consumption is described. The program was designed with the characterization of the Portuguese situation in mind, but the same approach is easily adaptable to other countries. The program uses a specific consumption data analysis methodology for the build-up of a database of demand patterns. The paper discusses the consumption analysis performed, as well as the development of the set of potential explanatory factors – including the key technical drivers related to the network characteristics, billing and customer statistics and a range of social-demographic variables – and illustrates with results obtained so far. The following points summarize the current stage:

- A protocol for data collection, with the aim of characterizing consumption at metering district level has been established.
- A mechanism for integration of data from different sources (i.e., geographical data, billing and customer data and social-demographic data) on GIS has been implemented to support the analysis.
- A statistical approach has been applied for the analysis of daily consumption patterns.
- A preliminary set of statistical variables has been proposed to characterize daily consumption patterns.
- A set of 20 metering districts from 9 water utilities are being used to test-run the program; the tests carried out so far are encouraging.
- Water utilities take part in this initiative on a voluntary basis.
- Access to the global database is open to participants as a retribution to their contributions; the basis for access to third parties will be established in due time.
- The program involves a multi-disciplinary team with expertise in water engineering, sociology and geography/GIS.

7 ACKNOWLEDGEMENTS

The authors are grateful to all participants in the INSSAA program (INSSAA, 2006) for providing the data and the insights essential to the analysis. This research is part of *Project P17: Technological innovation for an efficient management of the urban water, wastewater and solid waste services*, Study *E44: Water consumption studies to support water supply systems management* of the 2005-2008 Strategic Research Plan of LNEC. It was partially funded by FCT - Fundação para a Ciência e Tecnologia (Portugal), through project POCI/ECM/60917/2004.

8 REFERENCES

1. Alegre, H. (1992) - "Decision support tools for technical management of water distribution systems" PhD thesis, IST, Lisbon, Portugal (*in Portuguese*).
2. Alegre, H., and Coelho, S. T. (1993) - "A methodology for the characterisation of water consumption." Integrated computer applications in water supply Research Studies Press Ld., UK, 369-384.
3. Alegre, H., Coelho, S. T., Almeida, M. C., and Vieira, P. (2005) - "Control of water losses in transmission and distribution systems", IRAR, LNEC and INAG, ISBN 972-99354-4-0, Lisbon (*in Portuguese*).
4. Alegre, H.; Baptista, J.M.; Cabrera JR., E., Cubillo, F.; Duarte, P.; Hirner, W.; Merkel, W.; Parena, R. (2006) – "Performance indicators for water supply services", Second Edition, Manual of Best Practice Series, IWA Publishing, London, ISBN: 1843390515.
5. Burnell, D. (2003) - "Lifestyle and its effects on domestic use." *Advances in Water Supply Management*, C. Maksimovic, D. Butler e F. A. Memon, A. A. Balkema, Imperial College, London, UK, 647-656.
6. Census (2000) - "Census of Population and Housing." U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, Washington.
7. Census (2001) - "Census 2001: XIV Population Census of Portugal and IV Housing Census of Portugal." National Institute of Statistics, Lisbon.
8. Coelho, S. T. (1988) - "A System for Demand Analysis and Forecasting in Water Supply Systems," Master Thesis, University of Newcastle Upon Tyne, Newcastle, UK.
9. De Marinis, G., Gargano, R., and Tricarico, C. (2006) - "Water demand models for a small number of users." *8th Annual Water Distribution Systems Analysis Symposium*, Cincinnati.
10. Garcia, V. J., Cabrera, E., Garcia-Serra, J., Arregui, F., and Almandoz, J. (2003) - "Stochastic prediction of the minimum night flow demand in a district metered area." *Advances in Water Supply Management*, C. Maksimovic, D. Butler e F. A. Memon, A. A. Balkema, Imperial College, London, UK, 665-672.
11. INSSAA (2006) - "National Initiative for Water Supply Network Modeling (www.dha.lnec.pt/nes/iniciativa)." LNEC, Lisbon (in Portuguese).
12. IWA (2000) - "Blue pages on losses from water supply systems", International Water Association (IWA).
13. WRC (1994) - "Report f: Using night flow data." Technical report, UK Water Industry. Engineering and Operation Committee.