Analysis of household night-time consumption

D. Loureiro, R. Borba, M. Rebelo, H. Alegre, S. T. Coelho
National Civil Engineering Laboratory, Av. Brasil, 101, 1700-066, Lisbon (Portugal)

D. I. C. Covas, C. Amado, A. Pacheco
Instituto Superior Técnico - Technical University of Lisbon, CEHIDRO - Department of Civil Engineering, CEMAT - Department of Mathematics, Av. Rovisco Pais, 1049-001, Lisbon (Portugal)

A. Pina
SMAS Oeiras-Amadora Water Utility, Oeiras (Portugal)

ABSTRACT: Estimation of household consumption is one of the key elements in the analysis of water losses in distribution networks. Household night consumption is very specific of each network and socio-demographic context, and can vary according to multiple factors. This paper describes an on-going study aimed at assisting a Portuguese water utility in the development and implementation of a methodology to quantify household night consumption based on expedite rules, dependent on socio-demographic and billing factors. A stochastic model was developed to simulate household night consumption and was tested in 8 study areas. The paper describes the selection of case studies, the collection and processing of field data, the establishment of significant night consumption metering periods based on cluster analysis, the identification and characterization of rectangular pulse events, and the process of identification of and correlation with possible explanatory factors

1 INTRODUCTION

The assessment of water losses in water distribution systems, in particular the analysis of real (or physical) water losses, involves measuring minimum night flows and their desegregation into household consumption, non-household consumption, exceptional consumption and real losses.

Previous studies have concluded that rules to estimate household night consumption are specific of each water utility, since they can be rather dependent on socio-demographic factors (e.g., property type, household size, daily habits, etc.), technical factors (e.g., pressure) and climatic factors (UK Water Industry 1994; Warren 2002). Many publicly available results of household night consumption are connected to data obtained in the United Kingdom (UK Water Industry 1994). These results tend to be adopted elsewhere, as is the case in Portugal, due to the local unavailability of similar studies. However, aspects such as the inexistence of generalized household water metering, the predominance of detached houses for living and a generalised use of domestic tanks are major constraints for the extrapolation of British parameters to other regions with different contexts.

At the household level, several authors describe consumption as being based on rectangular pulse events characterised by three variables: arrival time over day, duration and intensity (Buchberger and Wells 1996; Garcia 2003; Garcia *et al*. 2003). Intensity and duration can be described in terms of adequate statistical functions, while the time of occurrence of a pulse can be described as an independent stochastic process (Buchberger *et al*. 2003; Garcia 2003).

A Poisson process enables possible inference for a larger number of users (spatial scale) and larger time steps (time scale) (Bosq and Nguyen 1996). This possibility is very important since it is not feasible to have individual detailed records for every household of a district metering area (DMA) and statistical inference can be a good alternative to estimate household night consumption.

Previous work indicated that household night consumption is influenced by socio-demographic variables such as number of occupants and type of dwelling (Warren 2002).

This paper describes a on-going study aimed at assisting a Portuguese water utility in the development and implementation of a methodology to quantify household night consumption based on expedite rules, dependent on socio-demographic and billing factors. The focus of the paper is the methodology developed and adopted, which significantly differs from the approach used in the UK. The main steps of each module of this methodology are described: 1) area selection, 2) data collection, validation and processing and 3) consumption modelling and presents the stochastic model developed. The application to case studies of the above-mentioned utility is presented.

2 METHODOLOGY

The proposed methodology for household night consumption estimation (Figure 1) is based on a general methodology developed by Loureiro (2006) for analysis of consumption. The same methodology is presented in a companion paper - Loureiro *et al.* (2009) – though applied at a different space scale and with a different objective (i.e., to characterize daily consumption patterns).

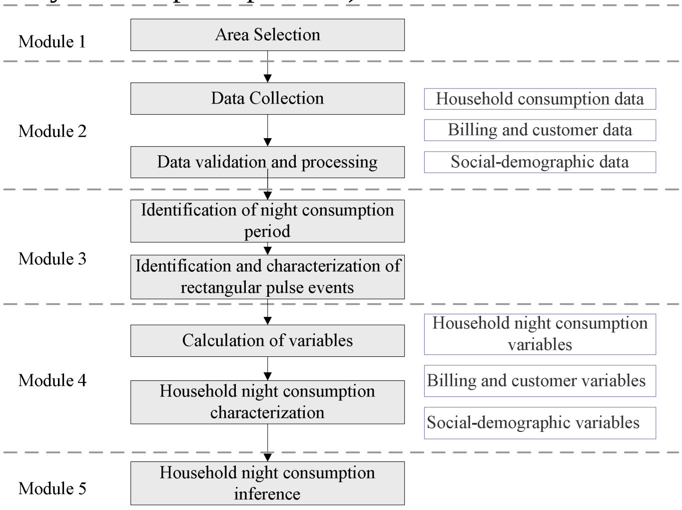


Figure 1 – General methodology.

Module 1 includes the selection and boundary definition the study areas to be monitored and the establishment of telemetry requirements (i.e., consumption meter, pulse emitter, datalogger, and communication system). An accurate identification of consumption is necessary to separate leaks from consumption. For this reason study areas should correspond to small units (~ 30 users), a reduced time step (e.g., 1 minute) and very accurate consumption meters on minimum flows are required. The data collection programme should have duration compatible with the analysis of consumption during winter and summer periods (i.e., higher than 3 months).

The objective of Module 2 is to collect the data. It requires preliminary tests to verify that consumption and pressure data are being adequately collected. Particularly, testing data collected should be analysed to validate telemetry systems installed and understand how the consumption meter is working. Telemetry data must always be validated with local readings. In this case study, data about billing and customers data was provided by the water utility and socio-demographic data was obtained from the last Census.

The goal of Module 3 is to develop a stochastic model for household night consumption. This model allows an analytical description of household night consumption and statistical inference of consumption for different time and spatial scale units. Previously the identification of minimum night consumption period was performed for each study area. A consumption model was developed based on approaches already published (Buchberger and Wells 1996; Garcia 2003) and adapted for the case studies considered.

The objective of Module 4 is the identification of the most important factors influencing household night consumption. Billing, customer and socio-demographic variables were established. Exploratory data techniques were used to identify inter relations.

Module 5 focuses on the extrapolation of household consumption to DMA level based on consumption parameters obtained for study areas (Module 3). Thus it is necessary to know the number of users and the time used to access water losses.

3 CASE STUDIES SELECTION

Case studies selection involved the identification of samples of household users (i.e., number and type of household users) from existing DMA and definition of meter and telemetry requirements to achieve study objectives.

Household night consumption should correspond exclusively to residential users and monitored as clusters of similar characteristics, located in small areas (e.g., building or a set of buildings at street level) and not at household level. This type of study area does not require the replacement of existing household water meters, or the involvement of clients. In terms of telemetry system, this option is less demanding in terms of data logging and communication systems. It has the disadvantage of measuring also possible water losses (i.e., at pipe, service connection and private connection pipe level) besides household consumption (i.e., use and plumbing losses).

In order to obtain study areas with a homogeneous number of household users, excluding non-household users and with a reduced water losses component, study areas with approximately 30 household users were selected (Figure 2). Study areas with this size allow a better separation between consumption and water losses and a proper modelling of household night consumption.

Figure 2. Study area map and meter location.

Each study area was selected in order to have socio-demographic (i.e., building, population, family) and billing characteristics (i.e., per household consumption and tariff structure variables) similar to the DMA where they are included. Previously, a characterization of consumption in each DMA has been performed and relations with socio-demographic variables have been found (Loureiro *et al.* 2009). A set of eight study areas were select from 8 DMA (Table 1).

Table 1. Study areas characteristics.

Study area	Consumption meter location	Number of users	Consumption per client (lclient/day)
ALF-sc	Service connection (sc)	21	290.0
AA-s	Street (s)	31	228.6
BA-s	Street (s)	28	270.2
FC-s	Street (s)	23	285.5
FN-s	Street (s)	32	198.7
NO-s	Street (s)	28	576.1
QMP-s	Street (s)	33	253.1
SA-sc	Service connection (sc)	20	282.9

Due to study requirements, in terms of high accuracy on household night consumption, and the uncertainty associated to maximum consumption values, a combined consumption meter has been selected (Figure 2b). This meter has a nominal diameter of 50 mm and extreme flows of Q_{min} = 6 l/h and Q_{max} = 90 m^3/h, in order to cover a large range of flow values. Although with an initial cost significantly higher when compared to a simple traditional consumption meter, this type of meter allows the elimination of pressure drop problems and minimizes the degradation of consumption meters to maximum consumption values, which can result in lower O&M costs. At the same time, the combined meter allows satisfying consumption meter requirements in terms of household night consumption. This type of meter has also the advantage of being easily applicable in other study areas with higher number of clients.

In terms of pulse emitter type, the main requirements were a high resolution and the minimization of electromagnetic interferences. Thus, an optical solution has been selected and installed, allowing a resolution of 1 pulse/litre and 1 pulse/ 0.1 litres in the principal and secondary consumption meter, respectively. In terms of data logger, the most relevant characteristics considered were the number of and type of input channels (i.e., digital and analogical), protection class, frequency of data reading and data transmission, type of data transmission (i.e., local or remote) and battery life.

In this study, a time step of 1 minute was adopted to collect consumption (i.e., 2 digital channels) and pressure (i.e., 1 analogical channel) during 12 months. Loggers IP 68 (i.e., protection class), with capacity to collect data during 15 days, with batteries with 6 month of duration were selected. Data transmission can be local, using a cable or remote, using GSM. Due to the volume of data generated per day and time necessary to communicate, data was sent daily to a central. Evaluating network communication strength during a preliminary campaign is fundamental. Data consumption and pressure is accessed via web.

a) b)

Figure 3. Equipment installed during a) preliminary data testing programme and during b) the collection programme

4 DATA COLLECTION AND PROCESSING

Consumption and pressure data were collected and processed in order to validate telemetry systems installed and to understand how the consumption meter is working. Telemetry data was compared with local readings. Billing and customer data were provided by the water utility and socio-demographic data were obtained from the last Census carried out in 2001.

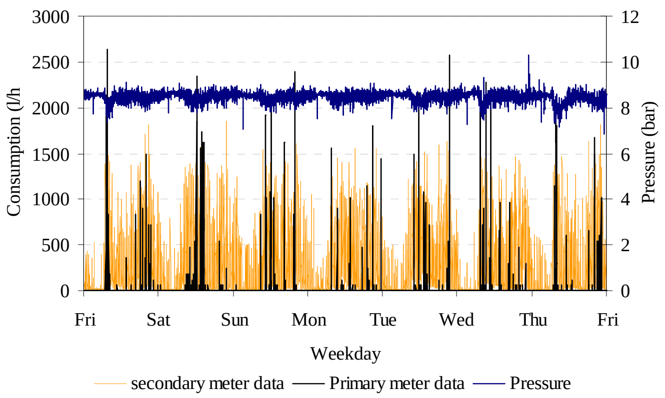


Figure 4. Consumption and pressure data collected at study area “SA-sc”.

Figure 4 shows consumption and pressure data collected at study area “SA-sc”. Consumption data was collected from two channels, since the combined meter is composed by two meters: i) secondary meter (volumetric meter) with high accuracy on minimum consumption values; and ii) primary meter (Woltmann meter) to measure maximum consumption values.

5 ESTABLISHING A NIGHT CONSUMPTION METERING PERIOD

Previously to consumption modelling, it is important to establish a homogeneous minimum night consumption period in each study area. In addition, it is important to detect week days that do not follow the same behaviour during the night. To achieve these objectives, clusters analysis has been used. Agglomerative hierarchical clustering has been adopted and different distance measures and types of linkage were tested. Figures 3 and 4 show the results of agglomerative hierarchical clustering using Euclidean distances and Ward's method as type of linkage, for study areas SA-sc and QMP-s.

In the case of SA-sc, three clusters were obtained, grouping hour consumption respective to 3-4 a.m., 2-5 a.m. and isolating consumption at 1 a.m. According to Figure 5, the period 3-4 a.m. corresponds to the minimum night consumption period.

In the case of QMP-s also three grouping clusters ware obtained: i) 1 a.m.; ii) 2-3 a.m.; iii) 4-5 a.m. (Figure 4) and the minimum night period corresponds to 4-5 a.m (Figure 5). These periods were adopted for household night consumption modelling.

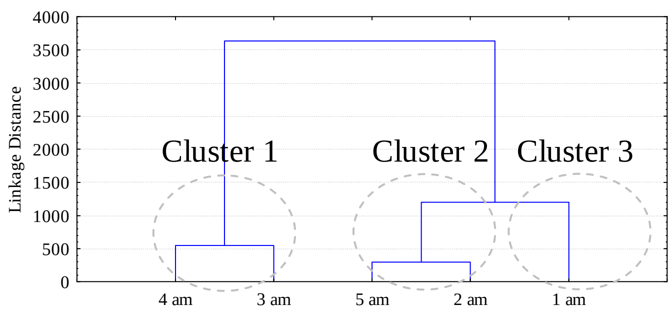


Figure 5. Dendrogram for variable "hour" during night period for study area "SA-sc".

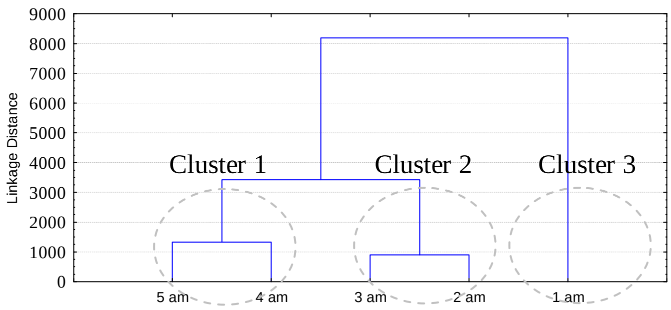
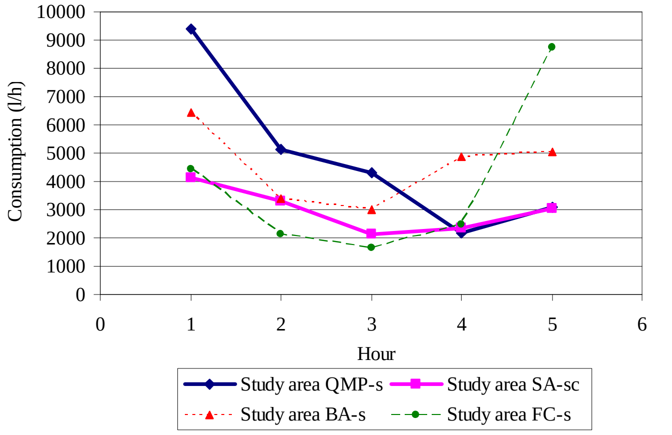


Figure 6. Dendrogram for variable "hour" during night period for study area "QMP-s"

In both study areas household consumption at 1 a.m. is completely different from the remaining hours analyzed. This type of analysis allows also the identification of seasonal effects on household night consumption.



.

Figure 5. Hour night consumption for study areas QMP-s, SA-sc, BA-s and FC-s.

6 IDENTIFICATION AND CHARACTERIZATION OF RECTANGULAR PULSE EVENTS

Since in study areas night consumption is characterized by a low rate of events and most of them do not overlap, consumption behaviour can be considered similar to a single user. Converting consumption into single equivalent events (SERP), Buchberger and Wells (1996) found that the occurrence of pulses can be described as an independent stochastic process. This result was later confirmed (Garcia 2003). Poisson process allows extrapolating the results obtained for a larger number of users (i.e., spa-tial scale) and larger time steps (i.e., time scale) (Bosq and Nguyen 1996). For these reasons and considering the objectives of the study and algorithm to obtain and characterize SERP was developed based on these assumptions.

Prior to SERP calculation, plumbing losses were separated from consumption. In general, minimum night consumption in the study areas ranges between 6 l/h and 40 l/h. Figures 6 and 7 show daily variation of minimum and average night flow and pressure for study areas "SA-sc" and BA-s.

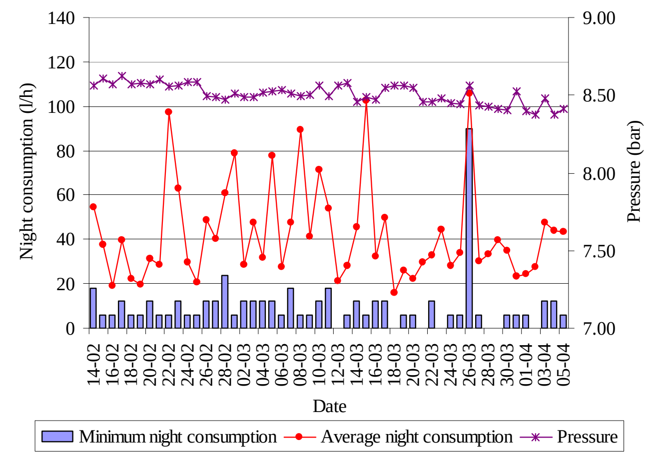


Figure 6. Minimum night consumption, average night consumption (2:00-5:00) and pressure for study area "SA-sc".

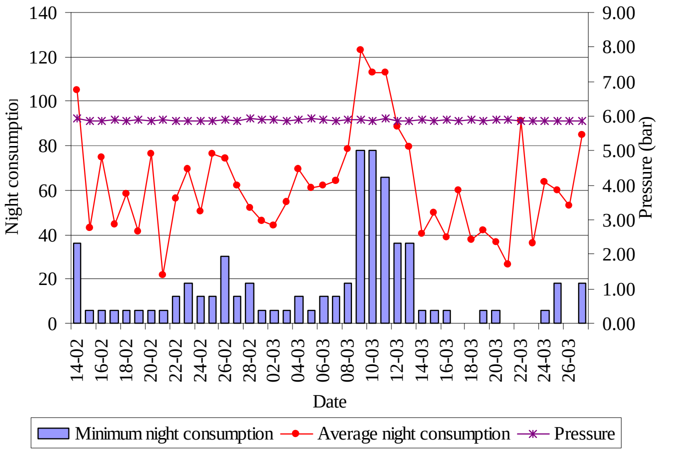


Figure 7. Minimum night consumption, average night consumption (2:00-5:00) and pressure for study area “BA-s”.

The algorithm for calculation of SERP is composed of two steps. The first step, designated by smoothing process (Buchberger and Wells 1996) consists in obtaining blocks of events (**Error! Reference source not found.**).

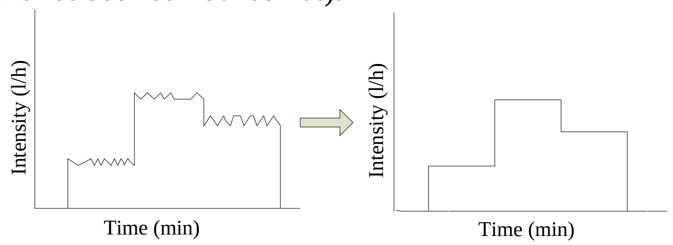


Figure 8. Signal smoothing.

The following step is the pulse separation of coincident water use for which three basic underlying assumptions were used: (i) two pulses can neither start nor end at the same time; (ii) coincident water use causes an increase of intensity; and (iii) at most two events occur simultaneously. The latter assumption, which is the less evident, was tested and confirmed for the study areas analysed.

In overlapping events, three situations can occur: (a) two pulses partially overlap; (b) a pulse completely contains another; or (c) three pulses occur sequentially without overlapping and in all situations there are 3 pulses with initial time t_i, duration d_i and intensity I_i. The first situation (Figure 9) is detected when the intensity of the middle block I_2, is between 85% and 100% of the sum of the other two (Garcia 2003), in which case the three pulses are separated into two pulses, the first with initial instant t_1, duration d_1+d_2 and intensity I_1* and the second with initial instant t_2, duration d_2+d_3 and intensity I_3*, where I_1* and I_3* are respectively I_1 and I_3 minus a correcting factor so that the volume is preserved. In the second situation (Figure 10) the intensities of the first and last block are similar. In this case the three uses are separated in two SERPs as follows: the initial block of the first one is t_1, its duration is $d_1+d_2+d_3$ and its intensity a weighted average of I_1 and I_3; as for the second SERP, the initial instant is t_2, the duration is d_2 and the intensity is I_2 minus the weighted average calculated above. If the conditions for either case aren’t met then all three uses are considered to be isolated events.

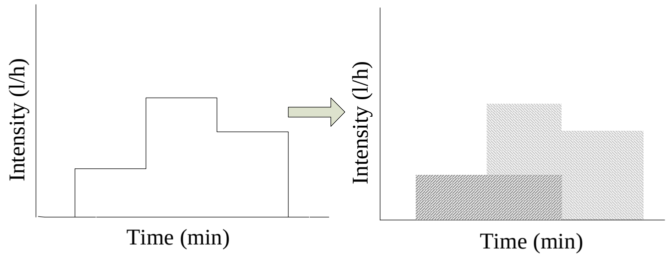


Figure 9. Pulse separation (case a).

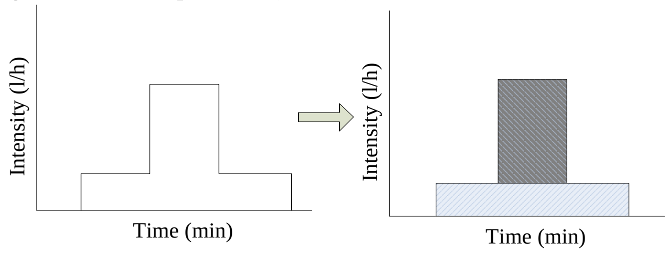


Figure 10. Pulse separation (case b).

The number of households occupied at a given time follows a Poisson distribution and the intensity follows a Weibull Distribution, in agreement with previous studies (Buchberger and Wells 1996; Garcia 2003).

Figure 11 illustrates the adjustment of busy homes that is achieved by the Poisson distribution, where the red line represents the expected frequencies of “Busy Homes” and the striped area represents their absolute frequencies from the study area FC-s. For the Chi-Square test the p-value is approximately 0.07.

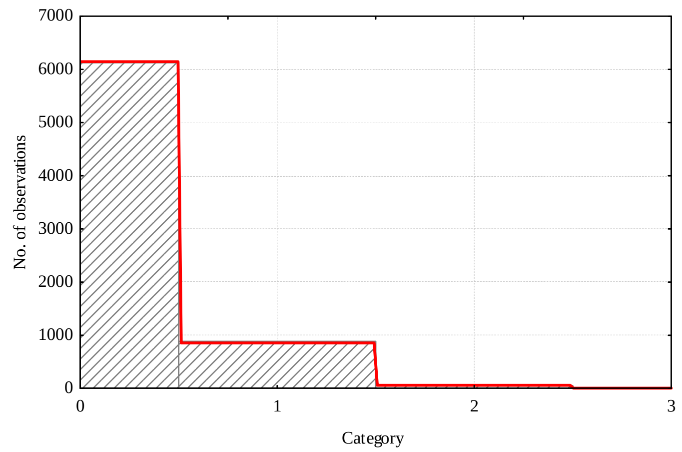
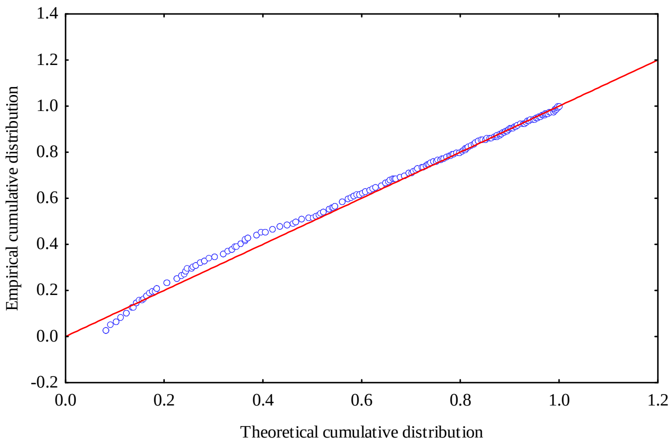


Figure 11. Absolute frequencies of “busy Homes” data and their expected values under a Poisson model.

Figure 12 presents the PP Plot adjusting the observed intensities from the study area FC-s to a Weibull distribution and again the observed data fits this distribution reasonably well.



.

Figure 12. Weibull PP-plot for “Intensities” data.

7 CALCULATION OF VARIABLES

Social and demographic variables considered were obtained according to the same methodology used by Loureiro *et al.* (2009). Census information (2001) was obtained considering the number of service connections monitored in each study area as regarding to the total number of service connections in the respective sub-section unit provided by Census.

According to Table 2, study areas can be grouped according to the analyzed socio-demographic characteristics. In this sense, FC-s and AA-s are urban areas with a predominance of elderly population living in small dimension families with low social mobility and living in old 2-floor buildings, whereas another cluster opposes BA-s, ALF-sc and SA-sc, where inner population can be described as younger people living in more recent buildings, integrating families with children and adolescents but sharing a low social mobility with the former cluster of DMA. The remaining sectors, namely FN-s, NO-s and QMP-s are more heterogeneous in terms if social and demographic features but share a high level of families with elderly elements and a higher social mobility.

Table 2. Study area socio-demographic characterization.

Area	Buildings (2 floors)	Families - adolescents	Families - elderly	Social mobility
ALF-sc	00.00	32.63	12.50	17.55
BA-s	65.78	44.28	5.71	3.86
FC-s	62.14	17.09	36.75	25.33
FM-s	00.00	00.00	44.44	34.48
NO-s	40.38	19.67	45.90	37.89
QMP-s	00.00	15.21	44.56	29.24
SA-sc	00.00	42.39	10.86	18.82
AA-s	00.00	13.20	36.79	11.71

8 CONCLUSIONS

An on-going study is described, aimed at assisting a Portuguese water utility in the development and implementation of a methodology to quantify household night consumption based on expedite rules, dependent on socio-demographic and billing factors. It was possible to establish a methodology for the assessment of household night consumption based on a set of study areas representative of household users in DMA of water utility. Requirements for study area and telemetry systems have been established. A stochastic model for household night consumption has been developed and is being validated and consolidated for 8 study areas. Modules 4 and 5 are in development in order to: i) find the most relevant factors that influence household night consumption and; ii) extrapolate consumption for a different number of user and to a time step compatible with the water losses methodology implemented in the water utility.

ACKNOWLEDGEMENTS

The authors wish to acknowledge SMAS Oeiras & Amadora for providing the data and the insights essential to the analysis, and for funding the project. Dália Loureiro is grateful to the Portuguese Foundation for Science and Technology (FCT) for the PhD scholarship (SFRH/BD/23819/2005).

REFERENCES

Bosq, D., and Nguyen, H. T. (1996). *A course in stochastic processes. Stochastic models and statistical inference*, Kluwer Academic Publishers Group, Dordrecht.

Buchberger, S., and Wells, G. (1996). "Intensity, duration and frequency of residential water demands." *Journal of Water Resources Planning and Management*, 122(1), 11-19.

Buchberger, S. G., Carter, J. T., Lee, Y., and Schade, T. G. (2003). "Random Demands, travel times and water quality in deadends." University of Cincinnati, Cincinnati.

Garcia, V. J. (2003). "Modelación de la demanda urbana de água," Univ. Potécnica de Valencia, Valencia, Espanha.

Garcia, V. J., Cabrera, E., Garcia-Serra, J., Arregui, F., and Almandoz, J. (Year). "Stochastic prediction of the minimum night flow demand in a district metered area." *Advances in Water Supply Management*, A. A. Balkema, Imperial College, Londres, Reino Unido, 665-672.

Loureiro, D. (2006). “Consumption analysis methodologies for the efficient management of water distribution systems, PhD thesis program, Instituto Superior Técnico, Lisbon Portugal (in Portuguese).

Loureiro, D., Coelho, S. T., Ramalho, P., Alegre, H., Covas, D. I.C, Pina, A. (2009). “Correlating water consumption behaviour with billing, infra-structure and socio-demographic factors”, to be presented in CCWI 2009,

UK Water Industry. (1994). "Report f: Using night flow data.", UK Water Industry Research, London.

Warren, R. (2002). "Household night consumption." UK Water Industry Research, London.