MODELLING METHODS TO SUPPORT URBAN SEWER SYSTEM MANAGEMENT

By

Juan Pablo Rodríguez Sánchez

A thesis submitted for the degree of Doctor of Philosophy from Imperial College London

Environmental and Water Resource Engineering Section

Department of Civil and Environmental Engineering

Imperial College London

London, July 2012

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In the context of developing countries, problems such as unplanned and unregulated urban developments, poor water quality in the water courses, lack of sanitation coverage and water treatment facilities, lack of institutional co-ordination, mismanagement of water resources, financial constraints, lack of wise expenditure on the required infrastructure, and a conventional fragmented wastewater management approach pose particular challenges. A sensible way forward in developing countries, which are initiating the building of treatment infrastructure, is to start integrated analysis and more efficient planning control measurements at an early stage. It is not feasible to build costly wastewater treatment plants which are not properly integrated to the sewer system and which are not solving the river pollution problems.

As a part of integrated analysis, models can be used to gain better understanding of certain phenomena and to predict the spatial and temporal evolution of a system when looking towards an integrated management of urban drainage systems. Particularly, a key requirement for integrated sewer system models is data to quantify flows and pollution loads, ideally including uncertainty estimates. Because well-gauged systems are rare, particularly for large scale developing urban centres, there is a need for methods of synthesising these data sets. Furthermore, large urban centres also need appropriate modelling frameworks to support an optimised sewer system preventive maintenance scheme. The final goal is to determine a plan of action that achieves the best balance between proactive and reactive maintenance to minimize the overall cost of system operation.

With this context as background and using Bogotá (Colombia) as main case study, this Thesis aims: (*a*) at an improved understanding of pollutant load contribution at the sewer sub-catchment level under dry weather conditions to be used to support the development of an appropriate modelling framework to assess and improve the interaction of the operation of the sewer system and the treatment facility; (*b*) to propose and critically assess a novel statistical model, supported by an

exceptionally long and spatially detailed customer complaints data-base, to support proactive maintenance of the sewer system.

This Thesis demonstrates the potential for using mixed deterministic/stochastic models to characterise ungauged sub-catchment outflows and water quality during dry weather. The novel contributions are a new empirical model for gauged and ungauged catchment applications, and the simultaneous inclusion of autocorrelations and cross-correlations for water quantity and quality determinants in the error model. Such models, if used to generate inputs to parsimonious sewer system models, can be useful to quantify discharges of untreated sewage into receiving watercourses, and their uncertainty, thus allowing these sources to be included in detailed dynamic river water quality models. Besides this, the impact of source control measures within the system can also be estimated. This is important, as source control is an essential complement to centralised wastewater treatment and stormwater storage. Additionally, it is also possible to quantify the transfer of pollutant loads from the wastewater system into the stormwater system due to wrong connections, thus helping water utilities with the prioritisation of corrective interventions. Furthermore, model-based research, using scientifically defensible tools such as the ones presented in this Thesis, is central to evaluating the performance of management practices such as proactive maintenance, water demand management and greywater recycling in Bogotá.

Due to the wide scope of the aims of this Thesis, they reflect different but strongly interrelated areas in current need of further research. The contribution of this dissertation is, therefore, not only individual theses, but the recognition of the integrated application of advanced and state-of the art methods (i.e. models and analysis frameworks) to deal with different operational requirements in the context of large and complex sewer systems. More generally, the work also demonstrates the value of monitoring and modelling programmes, including having modellers actively involved in monitoring specification and operations.

DECLARATION

The work presented in this Thesis is my own except where otherwise acknowledged.

Juan Pablo Rodríguez Sánchez

ACKNOWLEDGEMENTS

Thank you first of all to my supervisors, Dr. Neil McIntyre and Prof. Čedo Maksimović, for the opportunity that they provided to conduct this research, their guidance, patience and support throughout my PhD. I am grateful for the critical, constructive and continued feedback that helped me to greatly improve this dissertation and my understanding of the context of my research. I would also like to sincerely thank Prof. Mario Díaz-Granados, without his support and guidance it would not have been possible to undertake this PhD. I am also grateful with my former undergrad and master's supervisor, Dr. Luis Alejandro Camacho. His passion for, knowledge on and commitment with the water resources research field inspired and encouraged me to pursue this PhD.

Special thanks also go to Colfuturo, Colciencias and the Univesidad de los Andes for funding my study. Thanks to the Empresa de Acueducto y Alcantarillado de Bogotá (EAAB), especially to Juan Carlos Penagos for allowing and providing most of the data used in this work. The support of Dr. Stefan Achleitner during my PhD study is also greatly acknowledged. Thanks to him: I had access to the City Drain toolbox used in Chapter 5 of this dissertation, the case study of Linz (Austria) was proposed as part of model development in Chapter 4 of this Thesis, and I had the great pleasure to visit one of the most influential research groups in the urban drainage field, the Prof. Wolfgang Rauch's group.

I am also deeply grateful to all my colleagues at the Environmental and Water Resource Engineering section, particularly to many friends I had the great pleasure to share the 411 (previously 411a) research room with. They provided me with a great day-to-day work environment and also with many memories of unforgettable, friendly and enjoyable gatherings outside the college. I want to specially thank to: Joao, Ioana, Angeliki, Amutha, Chen, Nuno, Li-Pen, Tingting, Art, Simon, Susana, Costas, Giovanny, Alexandra, Jennifer, Alberto, Rossana, Antonello, and Gani.

I would like to strongly thank my parents and my brother, Juan José, Leonor and Andrés Felipe. Only their love, faithfulness and example have permitted the effort that has resulted in this document. Lastly, and more importantly, I thank my beautiful partner (now future wife!), Naty. She has provided me with unconditional love and support, without which, this Thesis would not have been possible.

PREFACE

Four main Chapters in this dissertation (i.e. Chapters 2, 3, 4 and 5) are based on works by Juan Pablo Rodríguez. Although these works have co-authors, the substantial research and writing effort was that of the first author.

Chapter 2

Rodríguez, J. P., Díaz-Granados, M. A., Camacho, L.A., Raciny, I. C., Maksimović, Č. and McIntyre, N., 2008. Bogotá's urban drainage system: context, research activities and perspectives. Sustainable Hydrology for the 21st Century, Proceedings of the 10th British Hydrological Society (BHS) National Hydrology Symposium, Exeter, UK.

Chapter 3

Rodríguez, J. P., McIntyre, N., Díaz-Granados, M. A., Achleitner, S., Hochedlinger, M. and Maksimović, Č., (In review). A stochastic time series generator of dry weather loads to sewers. *Environmental Modelling & Software*.

Chapter 4

Rodríguez, J. P., McIntyre, N., Díaz-Granados, M. A., Quijano, J. P. and Maksimović, Č., (To be submitted). Monitoring and modelling to support wastewater system management in developing mega-cities.

Chapter 5

Rodríguez, J. P., McIntyre, N., Díaz-Granados, M. A. and Maksimović, Č, (2012). A database and model to support proactive management of sediment-related sewer blockages. *Water Research* 46(15), 4571-4586.

Chapter 2, which corresponds to the case study description, is also partially based on the following published manuscript that is co-authored by Juan Pablo Rodríguez:

Díaz-Granados, M. A., Rodríguez, J. P., Rodríguez, M. S., Penagos, J. C., Camacho, L. A., Achleitner, S. A., Maksimović, Č., and McIntyre, N., 2009. Towards a paradigm shift in urban drainage management and modelling in developing countries. *Revista de Ingeniería – Universidad de los Andes* 30, 133-150.

Juan Pablo Rodríguez led or substantially contributed to other manuscripts during the course of this PhD research but they are not included within this dissertation:

Manz, B. J., Rodríguez, J. P., Maksimović, Č. And McIntyre, N., 2012. Impact of rainfall temporal resolution on urban water quality modelling performance and uncertainties. 9th International Conference on Urban Drainage Modelling, Belgrade, Serbia.

Rodríguez, J. P., Achleitner, S., Möderl, M., Rauch, W., Maksimović, Č., McIntyre, N., Díaz-Granados, M. A. and Rodríguez, M. S., 2010. Sediment and pollutant load modelling using an integrated urban drainage modelling toolbox: an application of City Drain. *Water Science and Technology* 61(9), 2273-2282.

LIST OF ACRONYMS

ARIMA	Autoregressive Integrated Moving Average
BOD	Biochemical Oxygen Demand
CAR	Regional Environmental Agency of Cundinamarca (Corporación Autónoma Regional de Cundinamarca)
CCTV	Closed-Circuit Television
cdf	Cumulative Distribution Function
CIIA	Environmental Engineering Research Centre (Centro de Investigaciones en Ingeniería Ambiental)
COD	Chemical Oxygen Demand
CSO	Combined Sewer Overflow
CSODT	Combined Sewer Overflow Detention Tank
DNP	National Planning Department (Departamento Nacional de Planeación)
DO	Dissolved Oxygen
DWF	Dry Weather Flow
EAAB	Empresa de Acueducto y Alcantarillado de Bogotá
EPR	Evolutionary Polynomial Regression
EQS	Environmental Quality Standards
HPP	Homogenous Poisson Process
MAVDT	Ministry of Environment, Housing and Territorial Development (Ministerio del Medio Ambiente, Vivienda y Desarrollo Territorial)
NHPP	Non-homogenous Poisson Process

NSE	Nash-Sutcliffe Efficiency
pdf	Probability Density Function
RP	Renewal Process
RTC	Real Time Control
SDA	Secretaría Distrital de Ambiente
TSS	Total Suspended Solids
TTT	Total Time on Test Statistic
UV-VIS	Ultraviolet-Visible
WFD	Water Framework Directive
WWTP	Wastewater Treatment Plant

CHAPTER 1:
INTRODUCTION

The useful implementation of a holistically-oriented and model-based urban drainage system management approach, with particular focus on large urban centres in developing countries, motivated this Thesis. In this first Chapter, the increased interest in the potential benefits of an integrated urban water management approach is discussed. Challenges and opportunities for application of such holistic management to developing country cities are presented. Technical challenges arising from the identified research gaps are introduced. Finally, the overall aim and objectives of this Thesis are listed and the structure of the reminder of the dissertation is described.

1.1 Integrated urban drainage management

1.1.1 Generalities

Many countries are experiencing accelerated urbanization, a demographic trend which is particularly marked in developing countries and transitional economies (Maksimović and Tejada-Guibert 2001; Parkinson and Mark 2005; Szollosi-Nagy 2008). At the beginning of 20th century just 16 cities in the world, mainly in industrialised countries, had more than one million inhabitants. Now; almost 400 cities have more than one million people, and about 70% of these are located in developing countries (Cohen 2006). The growing concentration of people in urban areas and their associated activities are dramatically increasing the pressures on urban water systems. Such urbanization increases the risk of considerable impacts on the catchment water balance and water quality (e.g. floods, overflows, pollution of receiving waters) and consequently on ecological and chemical status of receiving water systems (Jacobson 2011).

As a consequence, there is considerable interest in the potential benefits of an integrated management approach for urban water systems. Under a holistic framework, the supply, wastewater, stormwater and groundwater systems are viewed as components of an integrated physical system which collectively provide water supply, wastewater sanitation, flood protection, surface and ground water management, and ecosystem maintenance and protection (Mitchell 2006; Fletcher et al. 2008). Urban wastewater systems consist of the sewer system, the waste water treatment plant (WWTP), and the receiving system (normally rivers). Most sewer systems and WWTPs have been designed, operated, and improved as separate entities as a consequence of the difference in their main functions. However, Rauch and Harremoes (1998) and Lau et al. (2002) concluded that, for example, combined sewer overflow (CSO) reduction is not necessarily the appropriate operational objective when aiming at pollution control. Assessing drainage systems on the basis of how often an uncontrolled emission event occurs does not always correspond to improvement in the receiving water quality (Beenen et al. 2011). These research findings motivated the recommendation to revise environmental regulations based only on the number of overflows per year with regard to the impact on water quality (Engelhard et al. 2008). The design of efficient technical measures for the reduction of water pollution requires that wastewater discharge regulations are driven by receiving water objectives (Rauch et al. 1998). The idea of such an integrated urban drainage management approach is not new. Beck (1976) discussed a “water quality system” which involves the water distribution network, the sewer system, the treatment plant and the river. However, it was not until the early 1990s that the concept of the holistic approach began to be adopted in academic studies (Mitchell et al. 2007). The Interurba’92 workshop was an initiative of the International Association on Water Quality to advance knowledge of the interactions between urban drainage systems and sewage treatment plants and of the influence of contaminants from different sources on receiving waters. A fundamental goal of this workshop was to increase awareness of the impact on the receiving water ecosystem (Lijklema et al. 1993). Currently, the potential benefits of a coupled operation of the sewer system and the WWTP, guided by what is best for the river water quality, are well recognized (Langeveld 2004). Each receiving water body exerts its own properties and must therefore be evaluated under a holistic framework with respect to the discharges from the urban catchment and the non-point sources of pollution (Solvi 2006; Beenen et al. 2011).

The effects of wastewater discharges and CSO spills into the receiving streams include depression of the dissolved oxygen (DO) balance, deposition and resuspension of suspended solid matter, accumulation of heavy metals and nutrients in deposited sediments, bacterial contamination and elevated sulphide concentrations (Beck et al. 1988). However, the quality of surface water is not only controlled by the input of pollutants. Natural phenomena, hydrologic regime, local morphological conditions, climatic and ecological characteristics of the river basin have important roles (Lijklema 1995; Schilling et al. 1997). In Table 1 the impacts on receiving water from separate and combined sewer systems are summarized, classifying the time scale of impacts as acute, delayed and accumulating.

Table 1. Impacts on receiving water from separate and combined systems (Adapted from Borchardt and Sperling (1997) and Schilling et al. (1997))

Time Scale	Characterization	Indicator Variable
Acute (over hours)	Hydraulic	Flow, shear stress, bed erosion
	Chemical	Toxic substances
	Physical	Suspended solids
	Bio-chemical	Oxygen depletion
	Hygienic	Bacteria, virus
	Aesthetic	Floating material, odour
Delayed (over days)	Hydraulic	Sediment carrying capacity
	Chemical	Toxic substances
	Bio-chemical	Oxygen depletion in the sediments
	Hygienic	Bacteria, virus
	Aesthetic	Floatables, debris, oil
Accumulating (over weeks/years)	Hydrologic	Flow regime, morphology
	Chemical	Persistent organic substances, heavy metals, formation of inorganic and organic sediments, nutrients
	Bio-chemical	Oxygen depletion (eutrophication)

In the last decades, changes in the approach of planning urban drainage occurred in most developed countries. New ways of assessing the performance of urban drainage systems are based on "*stream standards*" (or Environmental Quality Standards – EQS (Bauwens et al. 1996)) and no longer based on an "*emission standard*" (Pfister et al. 1998; Benedetti 2006). These changes offer an integrated vision of the urban drainage systems that serve as a basis in some environmental regulations such as the Water Framework Directive (WFD) in the European Union (Barth and Fawell 2001). Consequently, a paradigm shift in the definition of performance indicators for urban wastewater systems is being observed (Zabel et al. 2001; Butler and Schuetze 2005; De Toffol 2006; Engelhard 2006; Muschalla et al. 2009). The main aim is to quantify the efficiency of different measures in reducing the amount of pollutants discharged into receiving water bodies and minimise the consequent negative impacts on water quality. As a part of this analysis, models can be used to gain better understanding of certain phenomena and to predict the spatial and temporal evolution of a system when looking towards an integrated management of urban drainage systems. Simulation models play a crucial role in environmental management plans, because they can be used to apply the best available scientific knowledge to predict responses to changing controls, as stated by McIntyre (2004) regarding basin management plans.

Detailed model-based integrated studies of the sewer network – WWTP – receiving water system are relatively rare due to high complexity of the entire urban drainage system that prevents a simple connection of the existing detailed deterministic models of the individual subsystems (Rauch et al. 2002), and practical applications of the holistic approach are still limited. However, recent progress has been made in developing integrated modelling tools, allowing practical application (Willems 2008; Candela et al. 2009; Devesa et al. 2009; Freni et al. 2009; Freni et al. 2009; Freni and Mannina 2010; Schellart et al. 2010; Freni et al. 2011; Candela et al. 2012). Furthermore, in order to increase the application of such a holistic approach, the Central European Simulation Research Group prepared a guideline document proposing a generic procedure for integrated modelling (Muschalla et al. 2009). The guideline covers the aspects of system analysis, identification of relevant system, processes and evaluation criteria, model setup and analysis, calibration and validation, scenario analysis and documentation. From this guideline it is clear that it is not only important to identify possible causes of negative impacts and/or to determine the potential of the system to be optimised, but to identify the relevant state variables and significant processes (Maryns and Bauwens 1997; Meirlaen 2002; Huisman et al. 2003; Vanrolleghem et al. 2005). In order to include all these relevant aspects, interactions and components, different modelling approaches can be applied, where the main differences are the amount of data required, the information that can be obtained from the model, the analysis performed and the simulation period. The type of simulation required principally depends on the objectives of the modelling (Rauch et al. 2005). The degree of detail in the different elements must depend on the available data and the available knowledge of the processes which have to be modelled.

1.1.2 Challenges and opportunities in developing countries

In the context of developing countries, facts such as unplanned and unregulated urban developments, severe water quality problems in the water courses, lack of sanitation coverage and water treatment facilities, lack of institutional co-ordination, mismanagement of water resources, financial constraints, lack of wise expenditure on the required infrastructure, and a conventional fragmented wastewater management approach pose particular challenges (Lee 2000; Maksimović and Tejada-Guibert 2001; Parkinson and Mark 2005). Unfortunately in such conditions, the less affluent sections of the society tend to be the most vulnerable due to the weak regulatory capacity of authorities and the ineffective solutions often applied (Parkinson and Mark 2005; Jamwal et al. 2008). A sensible way forward in developing countries, which are initiating building of the treatment infrastructure, is to start integrated analysis and more efficient planning control measurements at an early stage (Vollertsen et al. 2002). It is not feasible to build costly WWTPs which are not properly integrated to the sewer system and which are not solving the river pollution problems (Camacho et al. 2002; Camacho and Díaz-Granados 2003).

Large urban centres in developing countries are in urgent need of the development and implementation of measurement programs and modelling tools at different levels of detail, considering overall urban water fluxes and various sanitation schemes. A consensus between relevant stakeholders has to emerge towards planning and implementing strategies aiming to improve the sanitation situation, taking into account technological, environmental and socio-economic considerations. Bogotá city is a prime example of a large urban centre under the previously described stresses. A considerable investment in the Bogotá urban drainage system is expected in the near to medium term. With the help of a specific investment loan from the World Bank, nearly US$500 million are going to be invested between 2011 and 2016 in the upgrading and expansion of the treatment facilities, flood control and environmental restoration of the Bogotá River. The relatively undeveloped state of the current sanitation system provides an excellent opportunity for integrated planning and development. Now is the time to develop plans towards an efficient integrated system which maximises the benefits from the resources available.

1.2 Modelling tools to support management of sewer systems in large urban centres

1.2.1 Characterisation of dry weather loads to sewers

Integrated urban drainage models – needed to support holistic water management – tend to require large amounts of data both as model inputs and as observed responses for model calibration. However, most urban drainage studies suffer from lack of monitored data due to the high resource requirements of measurement campaigns (Freni et al. 2011). Furthermore, databases of water quality measurements generally come from sampling programmes which are fixed in frequency and location, without the main goal of registering the system's dynamic responses as is required for modelling. Even many full-scale WWTP simulation studies do not use realistic influent dynamics (Gernaey et al. 2011). Besides this, field measurements should be considered carefully as their value is affected not only by how well they represent the continuous system (i.e. number of measurement points, time and space scales, and heterogeneity of effluents) but also by uncertainty related to sensor accuracy and precision, and sampling techniques (Deletic and Fletcher 2008).

Although modelling-based performance assessment of urban drainage systems is subject to high uncertainty, it is common to neglect uncertainty in system planning, design, operation and management stages (Bertrand-Krajewski and Muste 2008). Recently, there has been increased research in quantifying and assessing uncertainties, both in quantity and quality variables (Willems and Berlamont 2002; Willems 2008; Freni et al. 2009; Freni et al. 2009; Freni and Mannina 2010; Schellart et al. 2010). General conclusions from these studies include: (*a*) the contribution to uncertainty from water quality sub-models can be an order of magnitude higher than that from flow sub-models (however, the contribution of flow sub-models is not negligible); (*b*) the accuracy of a sub-model is likely to be more important if it is nearer the downstream point of interest (e.g. outputs from the sewer system sub-model have a much larger control on the WWTP sub-model than on the receiving water sub-model); (*c*) integrated models of urban drainage systems tend to be highly parameterized and uncertainty propagates differently depending on sub-model complexity; (*d*) low data availability leads to larger model uncertainty. These insights can lead to more suitable model selection, improved monitoring design, and better informed decision-making (Freni and Mannina 2012; Lan-Anh et al. 2012).

An important component of uncertainty is that in the sewer system model inputs. In general, it is known that assuming accurate flow and water quality input data during model identification and testing can lead to large bias in the model (McIntyre and Wheater 2004). Furthermore, understanding and quantifying uncertainty can help in the model construction process (Friedler and Butler 1996; Metadier and Bertrand-Krajewski 2011). Arguably, the estimation of input uncertainty is a prerequisite to system model identification and subsequent uncertainty analysis (Deletic et al. 2012; Lin and Beck 2012). To contribute to this estimation, a generator of sub-catchment wastewater outflows, for use as inputs to dynamic simulations of the larger sewerage system, is developed and evaluated as part of the Thesis (see Chapter 4).

Besides the input uncertainty estimation issue, most urban drainage studies assume that the quality of wastewater is homogenous regardless of the catchment area and the pollutant load is correlated solely with the equivalent population (Gasperi et al. 2008). However, where concentrations of inflows to sewer systems are not homogenous over the system, and where the spatial location of the source matters in terms of what is delivered to the WWTP, a spatially distributed representation of DWF flows and pollutant generation from sub-catchments may be needed. Chapter 4 therefore also aims at such representation.

1.2.2 Wastewater flows and pollutant concentrations modelling in sewer systems

Models of water quantity and quality in the urban drainage system can be used to support: (*a*) performance assessment and definition of required upgrades to the existing urban drainage system and (*b*) optimal planning and design of new systems (Bauwens et al. 1996; Vollertsen et al. 2002; Benedetti et al. 2006; Gamerith et al. 2011). The components of an urban drainage system model may include a sewer system composed of a series of sub-catchments, trunk and interceptor sewers, a WWTP and a receiving water body as mentioned above. A key requirement for modelling all of these components is monitoring data - to quantify model boundary conditions (e.g wastewater pollution loads) and to provide observations of model outputs with which to guide model development and to calibrate and test the model (Price and Catterson 1997; Freni and Mannina 2012).

As indicated above, development and application of integrated quantity and quality urban drainage models has been recently reported for many case studies, mostly of relatively small urban areas in western Europe (Vanrolleghem et al. 2005; Solvi 2006; Willems 2008; Benedetti et al. 2009; Devesa et al. 2009; Freni et al. 2009; Freni et al. 2009; Andres-Domenech et al. 2010; Freni and Mannina 2010; Schellart et al. 2010; Freni et al. 2011; Todeschini et al. 2011; Lan-Anh et al. 2012; Prat et al. 2012). From review of these works, important factors limiting the successful (i.e. both justified by rigorous testing, and useful for the intended task) development of integrated models involving much larger sewer systems, may be proposed as follows:

- There is lack of monitoring programs covering mixtures of multiple non-residential water uses (e.g commercial, industrial and office-use) and to assess pollutant dynamic behaviour in the sewer system, with which to define inputs to the system, to inform conceptualisations of catchments, and to calibrate and test models.

- There has been limited development and testing of stochastic modelling tools for characterising uncertainty. This includes uncertainty in the loads to the system and uncertainty in the system itself.

- There is no clarity regarding appropriate model spatial and temporal resolution, and complexity of process representation.

These three challenges are reviewed and addressed as part of this Thesis in the context of dry weather conditions, with the general objective of proposing and critically assessing potential ways forward for large developing city applications (see Chapter 5).

1.2.3 Proactive management of sediment-related sewer blockages

Wastewater and stormwater carry a variety of solid particles which, when hydraulic conditions do not assure their transportation, form deposits. Solids accumulated in sewer systems can affect discharge capacity, increasing flood risk and frequency of overflow spills. Recent research has identified that drainage system conditions can make larger contributions to the risk of urban flooding than the occurrence of heavy storm events (Arthur et al. 2009; ten Veldhuis et al. 2009; Caradot et al. 2011). For example, in England and Wales about 75% (>23,000) of sewerage derived flooding incidents per year are due to blockages (Arthur et al. 2009). In Australia, blockages affect almost 70,000 properties across the country every year (Marlow et al. 2011). Besides this, water from sewer flooding incidents is likely to be contaminated thus posing potential health risk to citizens via waterborne pathogen exposure (ten Veldhuis et al. 2010); furthermore, the flushing of accumulated sewer sediment is one of the major sources of pollutants in urban wet-weather flow discharges (Verbanck et al. 1994; Rodriguez et al. 2010; Manz et al. 2012). For example, 65 out of 70 water utilities recently surveyed in the USA are using sediment control in order to protect or improve stormwater quality (Black&Veatch 2010).

Traditionally, municipal water utilities have addressed the maintenance and operation of sewer systems with a reactive approach, solving any problems only after they cause a failure. However, the cost of sewer failure (i.e. the cost of service disruptions, adverse publicity, and health and safety problems) can be significantly higher than the cost of implementing proactive maintenance. Some studies have identified regular sewer cleaning as a cost-effective way of dealing with flooding problems (Ashley et al. 2000; ten Veldhuis et al. 2009; Caradot et al. 2011; ten

Veldhuis and Clemens 2011), and a move away from reactive maintenance to a more proactive approach is encouraged (Fenner 2000). However, Wirahadikusumah et al. (2001) indicated that the major reason for using reactive approaches is lack of monitoring and record keeping, so that a system's deterioration is not evident until major failures occur. Lack of data on the condition of sewers also hinders the development of predictive models and the evaluation of effects of changes in the maintenance policy (Wirahadikusumah et al. 2001).

Efficient monitoring of sewer system condition would ideally include a survey of sediment accumulation as well as the structural condition of the sewer. There are a variety of technologies available for assessing sediment accumulation in sewer systems such as closed-circuit television (CCTV), and sonar/ultrasonic and laser profiling (Feeney et al. 2009). Furthermore, recent advances in acoustic-based instrumentation are allowing more rapid inspections (Bin Ali et al. 2011; Romanova et al. 2011). However, due to cost and time constraints, collection of data on sediment accumulation covering the entire sewer system is not sustainable in large urban areas (Fenner et al. 2000; Mashford et al. 2011). Therefore, analytical procedures are required for forecasting which of the sewer system structures are in most need of cleaning (Fenner and Sweeting 1999).

The complex physical mechanisms of sediment deposition, and the number of factors that may contribute to a sediment-related blockage, make forecasting a challenging task (Bachoc 1992; Laplace et al. 1992), and successful forecasting procedures are, in general, those based on statistical analysis of recorded sediment-related blockage events. Water utilities in many cities have dedicated call centres to handle customers reporting problems occurring in the urban drainage system. However, despite the availability of customer complaints and/or failure records there has been little formal analysis of the data, partly because of the lack of effective data storage (Fenner 2000; Arthur et al. 2009; ten Veldhuis et al. 2010). In cases where structured customer complaints/failure databases exist, these can be used to identify and study the various hydraulic and serviceability problems that are occurring within the sewer system (Arthur et al. 2008; ten Veldhuis et al. 2009; ten Veldhuis and Clemens 2010; ten Veldhuis et al. 2010; Caradot et al. 2011; ten

Veldhuis and Clemens 2011). In particular, Arthur et al. (2008) have combined complaints data and detailed hydraulic modelling to study blockage formation; and ten Veldhuis and Clemens (2010) and ten Veldhuis et al. (2010) have shown that automatic procedures to filter and classify complaints data can support prioritisation of operations. However, this previous research was applied to relatively small urban areas (ranging from about 20,000 to 170,000 inhabitants) and identified the lack of studies at much larger spatial scales. This Thesis describes an exceptionally long and spatially detailed data set from a customer complaints database covering the 7.5 million inhabitants of Bogotá which allowed implementation and testing of a new statistical model to help identify the factors most affecting sediment-related blockages, to identify how the blockages rates are distributed in time and thus to help prioritize sewerage maintenance (see Chapter 6).

1.3 Overall aim, hypotheses and objectives of this Thesis

The overall aim of this Thesis is to contribute towards improved model-based management of wastewater systems in large developing cities. To facilitate effective application of modelling frameworks and tools to large scale and complex sewer systems, this dissertation aims at studying different hypotheses that could help to guide model development:

- There is a needed for stochastically represent sources of variability when characterising sub-catchment wastewater outflows

- Observed variability of wastewater flows, pollutant concentrations and pollutant loads within sewer systems during dry weather conditions is largely controlled by the sub-catchment outflows rather than in-sewer process within the main sewer system

- Sediment-related sewer blockage is a highly complex process that cannot be synthesised deterministically due to lack of suitable data sets to support model development. Instead, statistical approaches can provide useful insights on blockage rates that constitute valuable information for defining maintenance strategies

As part of hypotheses testing, different objectives are to be addressed in this dissertation. In general, they aim at delivering demonstrated modelling tools and analysis frameworks for the sewer system, considering its interactions with treatment facilities and local receiving water courses. The main objectives can be summarised as follows:

- Develop an enhanced model to characterise pollutant load contribution at the sub-catchment level, both for gauged and ungauged areas, under dry weather conditions, thus providing tools to prioritise potential source control strategies.

- Identify a parsimonious model to assess and improve the interaction of the design and operation of the sewer system and the WWTP under dry weather conditions.

- Propose and assess a statistical model to support proactive cleaning of sediment-related sewer blockages.

1.4 Structure of the remainder of this Thesis

In *Chapter 2*, general considerations and discussion on previous research developments useful to give context and relevant to support the different theses contained in this dissertation are presented. Particularly, the Chapter reviews and discusses: (a) available modelling methods to characterise dry weather loads to sewers and their limitations for large scale modelling applications, (b) current limitations and unknowns regarding appropriate monitoring and modelling methods to support development of sewer system models for large urban centres, and (c) definition, triggering factors, available modelling approaches of sediment-related sewer blockages.

Chapter 3 presents a detailed overview of the main case study used in this Thesis: Bogotá (Colombia). The current status and development of the urban drainage system (for each of the main system components, i.e. sewer system – WWTP – receiving water courses), details on water resources management plans and available data sets are presented. This allows a better understanding about how this PhD research can contribute to this challenging case study.

Chapter 4 aims at an improved characterisation of pollutant load contribution at the sub-catchment level under dry weather conditions, thus providing some of the required inputs for large scale sewer system modelling. To contribute to this goal, a composite model is presented: (a) one sub-model that estimates mean wastewater flow and biochemical oxygen demand - BOD, chemical oxygen demand - COD and total suspended solids - TSS concentrations, (b) one that quantifies the mean transfer of flow and pollutant loads from the sanitary sewer system into the stormwater system via wrong connections, (c) one that represents mean infiltration flows into the sewer system, and (d) one that characterises flows and pollutant concentrations sub-daily variability around mean values.

Chapter 5 develops a suitable modelling framework to support operation and development of the wastewater system of Bogotá. Components of the framework covered here are: (*a*) the flow and water quality database (see Chapter 3), (*b*) a wastewater pollution load generator (see Chapter 4), and (*c*) a semi-distributed sewer network model, which aims at a complexity that matches the information available from the previous two components.

Chapter 6 proposes and critically assesses a novel statistical modelling supported by an exceptionally long and spatially detailed customer complaints data-base to support proactive maintenance of the sewer system. This Chapter illustrates the potential value of such databases and formal analysis frameworks for maintenance scheduling in large cities. Finally, *Chapter 7* critically discusses the work's significance and proposes the way forward.

CHAPTER 2:

LITERATURE REVIEW

In this Chapter, general considerations and discussion on previous research developments useful to give context and relevant to support the different theses contained in this dissertation are presented. This state-of-the-art literature review is sub-divided in the same way Thesis Chapters are defined. Chapters 3 and 4, and therefore their corresponding sections in this review, focus their attention on dry weather conditions.

2.1 Modelling methods to characterise dry weather loads to sewers

It is well known that the inflow to wastewater collection systems exhibits considerable variations in terms of both quantity and quality at different time scales (e.g. daily, weekly, seasonal, etc.). This dynamic behaviour results from, among other causes, the quasi-random usage of different appliances within different water use sectors, including residential, industrial, commercial, and office-use wastewater DWFs, as well as from infiltration and storm runoff. Although DWF from any one sector has a repetitive-like diurnal pattern at urban sub-catchment outfalls, it is still subject to random variability (Butler and Graham 1995).

Previously, different methodologies and models have been developed for generating DWFs and associated pollutographs. However, many of these methodologies are highly data-demanding in terms of surveys of domestic/industrial/commercial appliance usage (Butler and Graham 1995; Almeida et al. 1999). Various authors have attempted to overcome the need for appliance data by fitting Fourier series to observed DWF data (Carstensen et al. 1998; Bechmann et al. 1999; Langergraber et al. 2008; Mannina and Viviani 2009), while De Keyser et al. (2010) attempt to improve DWF estimation at ungauged sites by coupling normalised diurnal profiles with local population-equivalent data. Some studies have also recognised that there are sources of variability, such as varying frequency of water use and water consumption rates (Friedler and Butler 1996), that cannot be represented entirely deterministically and that adding a stochastic component to the model is beneficial (Carstensen et al. 1998; Bechmann et al. 1999; De Keyser et al. 2010). The stochastic term may include: the quasi-random nature of water use within each day (Butler and Graham 1995); variability at longer scales (e.g. seasonal effects or long-term trends); measurement error; and any other features of the observed diurnal pattern that cannot be explained by the deterministic model. Such mixed deterministic-stochastic generators can be used to model DWF profiles from gauged and ungauged sub-catchments including uncertainty. This is particularly valuable in large and complex urban areas such as developing mega-cities where, due to cost and operational constraints, comprehensive field collection of data for characterising DWF outflows from large numbers of sewer sub-catchments is not feasible.

In the early 70s, the concurrent availability of data sets (mainly characterising WWTP inflows) and an enhanced computational devices triggered the development of stochastic modelling of wastewater quantity and quality time series. Research interest was mainly focused on implementing models based on autoregressive integrated moving average (ARIMA) models for forecasting influent flows, and pollutant concentrations to WWTPs (McMichael and Vigani 1972; Berthouex et al. 1975; Barnes and Rowe 1978) but application of system identification analysis of single-input, single-output stochastic processes were also studied (Capodaglio et al. 1990; Zheng and Novotny 1991). But stochastic models have not only been used to represent WWTP inflows - they also have been developed to estimate pollutant load variations in sewers systems due to urban runoff wash-off processes (Harremoes 1988; Rossi et al. 2005) and use of household chemicals, pharmaceuticals and personal care products (Ort et al. 2005; Rieckermann et al. 2011).

2.2 Sewer system modelling in large urban centres

2.2.1 Monitoring to support sewer system models

Sewer system modelling in large urban centres ideally requires field data of flow and concentrations from the outlet of each sub-catchment. Different wastewater quantity and quality databases aiming at characterising sewer sub-catchment outflows have been reported (Wilkie et al. 1996; Pantsar-Kallio et al. 1999; Singh et al. 2005; Gasperi et al. 2008; Lamprea and Ruban 2011). These databases are in general very comprehensive in terms of the number of analysed wastewater quality determinants. However, they are of limited use for large scale modelling purposes: (*a*) they primarily characterised residential and commercial sub-catchments; (*b*) they were mainly focused on quantifying mean daily loads or only few sub-catchments were used to characterise sub-daily variability.

Most of reviewed monitoring programmes have been mainly based on manual or automated sampling. There are various known limitations of this method such as errors due to sampling conservation, transport and preparation, short duration campaigns, limited time intervals, and high proportional costs (Gruber et al. 2005). Despite these limitations, automatic samplers have been identified as a suitable mechanism in order to measure pollutant concentrations in sewer systems (Larrarte 2008). Nevertheless, nowadays by means of in-situ measurement technology it is possible to measure concentrations with a high temporal resolution, therefore allowing a better understanding of the sewer system behaviour. For example, ultraviolet-visible (UV-VIS) spectrometers allow in-situ real-time measurements of organic compounds (Gruber et al. 2004), with no sampling and no sample preparation (Gruber et al. 2005). A number of parameters can be measured simultaneously using only a single instrument. However, the application of UV-VIS sensor technology in sewer systems has been limited for many reasons such as difficult functioning conditions, lack of reliability, fast fouling and difficult access for maintenance (Hochedlinger et al. 2005). This raises the question of whether manual or automated sampling is suitable for characterising DWF dynamics within large sewer systems or if more advanced equipment, allowing a higher time resolution monitoring, are needed as well or instead.

2.2.2 Sewer system models

White-box, grey-box and black-box models

Model-based design, assessment or optimal control of wastewater systems generally correspond to cases where neither a “white-box” approach (i.e. the available knowledge of the physical properties of and the dynamic processes within a system is used, and none of the variability is treated as stochastic) nor a “black-box” approach (i.e. both the model structure and the parameterisation are derived and validated by statistical methods, implying that knowledge about the system’s physics is ignored) is ideal (Thordarson et al. ; Breinholt et al. 2011). It is recognised that including prior knowledge of the system is often necessary to construct adequately powerful models, but there is also need for representing additional sources of variability (Vollertsen et al. 2005; Vollertsen et al. 2005). For such cases a “grey-box” modelling could be useful, which in the context of this work is a deterministic model (either white-box or black-box based) with stochastic terms to account for uncertainties in model formulation and model inputs (Bechmann et al. 1999).

The grey-box approach requires decision about which properties of the system and which processes to explicitly represent in the model, which to neglect, and which to represent stochastically. The inclusion of processes that do not significantly contribute to the important outputs results in a model that is unnecessarily difficult to use and has unnecessary data demands (Almeida et al. 1999). Such decisions can be greatly assisted by considering the sensitivity of the required model outputs, in the context of this work the relevant WWTP influent data and untreated wastewater discharges (via CSOs) into receiving water courses, to these properties and processes. Previous research has demonstrated that, depending on WWTP configuration and characteristics, short-term fluctuations of a particular determinant may not be important (Langeveld et al. 2003). This, for example, can reduce the requirements of the quality modelling, since exact calculation of short-term dynamics is unnecessary. In some cases, depending on the relation of dissolved and attached fractions of a certain water quality determinant, biodegradation and/or sediment transport do not play an important role in the dynamic interactions between sewer systems and WWTPs. Furthermore, system sensitivity to flow and pollution load variations can determine if the model has to consider flow and concentration dynamics, or only flows with the assumption of constant concentrations (Langeveld et al. 2003).

Spatial resolution

Implementation of detailed sewer models at large spatial scales is currently limited due to data availability requirements, this being particularly marked if physically based distributed models are intended. To overcome such a limitation, Blumensaat et al. (2012) proposed an approach to generate close-to-reality sewer networks under data scarcity conditions thus potentially allowing detailed distributed hydraulic models to be applied to large systems. The approach combines the application of a surface flow accumulation algorithm to a digital elevation model with a routine for hydraulic network dimensioning (i.e. estimation of pipe diameter distributions). However, further development is still needed regarding hydraulic dimensioning and applicability to complex networks (i.e. those sewer systems comprising control structures and storage facilities). Besides this, Freni et al. (2008) demonstrated that using detailed distributed sewer system models under data scarcity conditions does not provide advantages over using simplified approaches for the purpose of providing inputs to detailed river water quality models. Consequently, it is proposed that a semi-distributed conceptual model can be used to represent the sewer system: In comparison with a distributed approach, this decrease supporting data requirements but still aims to provide a reasonable spatial representation of the sewer system and its spatial and temporal variability, thus enabling integrated assessment.

In-sewer processes

In many sewer systems, hydraulic retention provides sufficient time, of the same order of magnitude as in the WWTP, for the wastewater and solids to be transformed by in-sewer processes (Nielsen et al. 1992). The composition of wastewater with an age of only a few minutes or hours may be markedly different from wastewater which has been under transportation for many hours. This is mainly due to microbial growth and respiration in both bulk water and biofilms, solubilisation, enzymatic hydrolysis of macromolecules and hydraulic shear forces. Also sedimentation and resuspension may be important (Nielsen et al. 1992). Kaijun et al. (1995) carried a laboratory-scale study which aimed at understanding the nature of sewage changes during transportation and storage. They found that, for example at 30 °C, the COD removal capacity in an aerobic system was 130 mg (or 15 %) COD l^{-1} d^{-1} (about 5.5 mg COD l^{-1} h^{-1}) within the first three days. In the case of an anaerobic system the average COD removal capacity was only 38 mg (or 4.5 %) COD l^{-1} d^{-1} (about 1.6 mg COD l^{-1} h^{-1}).

Sewer systems can be considered to be reactors where physical, chemical and biological processes cause interactions between the aqueous, solid, and atmospheric phases. Processes in the water phase related to wastewater transformations are hydrolysis, reaeration and microbial processes. These processes may proceed under redox conditions determined by the availability of the electron acceptor (Hvitved-Jacobsen et al. 2002). The oxygen consumption and exchange processes in the water phase, in the biofilm and at the sediment/water and water/sewer atmosphere surfaces play an important role for changes in wastewater composition (Nielsen et al. 1992). Under aerobic conditions the main processes are growth of heterotrophic biomass, removal of readily available biodegradable organic fractions and hydrolysis of biodegradable substrate (Hvitved-Jacobsen 1998). The contribution of the attached biomass (biofilm) to the aerobic conversions is mainly determined by the DO concentration and the wetted surface area in relation to the volume. The contribution of the suspended biomass activity depends on the amount of active bacteria in the wastewater source, the state of the biofilm, the time of the day and specially the residence time (Huisman et al. 2004). Under anaerobic conditions the main processes involved are fermentation of readily biodegradable organic matter resulting in a production of volatile fatty acids, hydrolysis of biodegradable substrate, although at a much lower rate (i.e. 15%) than under aerobic conditions and sulphate reduction (Tanaka and Hvitved-Jacobsen 1998).

The importance of the processes for the sewer and the surroundings is not just caused by the removal and transformation of organic substrates – the electron donor – but is also a result of transformation of the electron acceptors exemplified by the formation of hydrogen sulphide from sulphate (Hvitved-Jacobsen et al. 2002). For example, Gudjonsson et al. (2002) encountered in field conditions that when wastewater temperature increased from 9°C to 14°C the electron acceptor conditions changed from being solely aerobic to being alternating aerobic-anaerobic. They concluded that at the highest temperatures the sewer was anaerobic for a significant part of the day. The latter means that without good and detailed data on the wastewater composition, the electron acceptor conditions can only be roughly estimated (Gudjonsson et al. 2002). In a sewer network, primarily aerobic and anaerobic conditions arise whereas anoxic conditions only exist if nitrate occurs in the wastewater (Hvitved-Jacobsen et al. 2002). The magnitude of the reaeration is a central process that determines if aerobic or anaerobic conditions exist. For low flux of oxygen into the wastewater (e.g. 3 to 1 g O_2 m^{-3} h^{-1} along the sewer line) – and a corresponding very low DO concentration – only an insignificant amount of COD is removed. Under such conditions, the distribution of the COD fractions is only slightly changed during transport in the sewer (Hvitved-Jacobsen et al. 2002).

The inclusion of water quality aspects and processes in sewer system modelling came together with the application of the integrated modelling approach. Most sewer models considered pollutants to be conservative until the middle of the 1980s. The conservative approach has been progressively replaced with one which takes into account transformation processes (Rauch et al. 2002). In the Interurba II conference, the concept of "*the sewer as a bioreactor*" was presented (Harremoës 2002). Regarding driving factors of in-sewer process, there are three variables that largely control the rate and nature of pollutant transformations (as reviewed above): (*a*) presence of DO (Hvitved-Jacobsen et al. 1998; Hvitved-Jacobsen et al. 2002), (*b*) bulk wastewater temperature (Gudjonsson et al. 2002) and (*c*) hydraulic retention time (Nielsen et al. 1992; Kaijun et al. 1995; Huisman et al. 2004). The potential importance of in-sewer transformations implies the need for detailed modelling of not just the main transformation processes themselves but these driving factors (Hvitved-Jacobsen et al. 1998; Almeida et al. 1999; Calabro et al. 2009); or at least for some simplified representation that represents the spatially and temporally averaged effects (Ahnert et al. 2005). However, there are some problems and difficulties to be faced in applying this concept, for example the limited amount of observations economically or logistically possible, thus the difficulty of quantifying process rates. Furthermore, detailed wastewater quality models are highly parameterised. Even though such complex models can be simplified through a sensitivity analysis (Calabro et al. 2009), sensitive model parameters still require considerable measuring efforts for their identification. Furthermore, if data constraints and/or system variability mean that uncertainty in pollution loads dominates the uncertainty in modelled WWTP influent, then there may be no real value in representing in-sewer processes (Flamink et al. 2005).

2.3 Sediment-related sewer blockage modelling

2.3.1 Hydraulic deterioration of sewer systems: definition and triggering factors

Hydraulic deterioration of sewer systems has been defined as a continuous process that reduces the discharge capacity through a reduction of cross sectional area and an increase in pipe roughness (Tran et al. 2010). It is caused, among other factors, by sediment accumulation (Fenner and Sweeting 1999; Kannapiran et al. 2007; Tran et al. 2010; Marlow et al. 2011), which Chapter 6 focuses on. Previous research has looked at gaining better understanding of the role of sewer network properties in promoting solid deposits and sewer blockages (Bachoc 1992; Laplace et al. 1992; Chebbo et al. 1995; Fenner and Sweeting 1999; Gerard and Chocat 1999; Savic et al. 2006; Arthur et al. 2008; Ugarelli et al. 2009; Ugarelli et al. 2010; Marlow et al. 2011). It is known that the propensity for sediment deposition depends upon the location of a sewer in a network and some of its physical characteristics. Nevertheless, there is no consensus on which physical properties can be considered as influential factors. Fenner and Sweeting (1999) have concluded that urban catchments with brick sewers, long pipe lengths, small diameters, shallow depths, moderate to slack gradients and foul sewers tend to have more frequent blockages. Ugarelli et al. (2009) concluded that, in addition to previous properties, system age and function (sewage, stormwater and combined sewer systems) seem to have a marked influence on proneness to blockages. However, Ugarelli et al. (2010) did not find the slope to be a clear significant factor, possibly due to limited data availability. Tran et al. (2007) argued that the number of nearby trees and climatic conditions can also influence the state of deterioration. Tran et al. (2010) identified that older pipes or those with more surrounding trees, tend to be in better hydraulic condition. In the same study, Tran et al. (2010) also concluded that soil type and soil wetness were not significant factors and no conclusions could be drawn about the significance of pipe slope and pipe depth. More recently, Marlow et al. (2011) carried out a web-based survey to collate expert opinion from 21 Australian water utilities on factors that influence sewer blockage rate (due to sediment build-up, physical objects within the sewer, sewer defects, accumulation of fats, oil and grease and tree roots). Their results indicate that utilities consider drought, sewer attributes, tree coverage, climate and tree planting policy as the most relevant factors. However, a detailed analysis using blockage data from two different water companies showed that the number of blockages was driven primarily by pipe age, pipe diameter and water consumption rates. Although the predisposition for sediment accumulation is strongly related to the system's physical characteristics, it also depends on the flow field characteristics, the nature of the particles and the concentration in suspension and/or near the bed (Ashley et al. 2000).

Even though many factors may be important, pipe size (diameter) and pipe age have most consistently been identified as significant factors affecting sewer system blockage likelihood. As stated by Arthur et al. (2008), these previous research results should help sewerage asset managers to focus proactive maintenance activities on identifying where additional future blockages may occur.

2.3.2 Available modelling approaches

Hydraulic deterioration models of sewer systems can be classified into two main groups: (*a*) physically-based and (*b*) statistical models. Physically-based models are developed based on the understanding of the physical mechanisms that govern sewer system deterioration. A key aspect of these models is that they can allow extrapolation of the underlying relationships from one system to another. However, limitations in process understanding and scarcity of data about local conditions lead to large uncertainty about the suitability of physically-based models (Fenner et al. 2007). Consequently, most deterioration models are of the statistical type based on observations of failure events.

With the aim of deriving probability distributions of time intervals between failures and identifying significant trends in the failure rates, Ascher and Hansen (1998) proposed a general procedure for analysing failure data from repairable systems. The framework consists of the following main steps: (*a*) validation of failure data (e.g. identification of outliers, errors or inconsistencies); (*b*) statistical analysis of number of cumulative failures versus time (e.g. to identify changes in failure patterns, effects of maintenance policy); and (*c*) description of failures with either a constant or time-varying failure rate using a homogenous Poisson process (HPP) model, more general renewal process (RP) models, or a non-homogenous Poisson process (NHPP) model (all described further later in Chapter 6). Jin and Mukherjee (2010) applied the Ascher and Hansen framework to predict the rate of blockages caused by pipe deterioration, variations in system usage load, construction defects and uneven soil support. They discussed the advantages of treating sewer blockages as a continuous stochastic process in which blockage events occur independently at a constant average rate. Using 10 years (1996-2006) of data from the City of Houghton (Michigan, US) (a small municipality with about 1,500 customers), Jin and Mukherjee (2010) concluded that all one-year subsets of blockage data (i.e. time between failures) were best characterised by an exponential distribution. However, subsets of two and three continuous years of data were generally best fitted by a Weibull or gamma distribution, with the exponential distribution being a close approximation. It was also identified that for even longer periods (i.e. 8 years) an exponentiated (three-parameter) Weibull distribution was needed to better represent observed heavy-tailed distributions. Jin and Mukherjee (2010) argued that the latter is mainly caused by the cumulative impact of more frequent blockages at certain times of the year (due to snow melt and time-concentrated increased human activity). They also demonstrated that HPP models are more applicable when the data are grouped into physically meaningful subsets, for example according to pipe diameter and the season.

Fenner and Sweeting (1999) and Fenner et al. (2000) presented one of the first attempts to explore the quality and extent of failure databases and discussed a number of ways in which the information may be analysed to help inform drainage engineers where best to target their efforts. They used 4 year long event databases to rank 48 grid squares (each about 0.25 km^2) into priority zones for action. An algorithm based only on the number of past events (totalling 256 for the studied urban area) in each square was found to be the best predictor of future blockage rates. The likelihood of future events occurring in a grid square was coupled with the impact of sewer blockage and a cost model, thus allowing prioritisation of maintenance between squares. In a second modelling stage, Bayesian analysis was used in an attempt to identify specific pipe lengths that should be targeted for maintenance. However, Fenner (2000) concluded that this was not possible without resource to some catchment-specific data analysis.

The evolutionary polynomial regression (EPR) method (Giustolisi and Savic 2006) was applied by Savic et al. (2006) for modelling blockage events using 10 year records from a sewer system in the UK. The study aimed to highlight pipe attributes that are the key variables in modelling blockage phenomena. The authors proposed a regression equation that expressed blockage rate as a function of diameter, equivalent slope, total length of pipes and number of pipes, which gave a coefficient of determination of 82%. A strong direct correlation with pipe diameter and an inverse correlation with total length were observed. The authors concluded that at sites where no records of blockage are kept, their regression models may be useful for predicting frequency of blockages. Although it was recognised that the significant explanatory variables may vary from one sewer system to the next, averaging EPR-based models over multiple sewer systems was found to achieve reasonable performance (Savic et al. 2009). (Ugarelli et al. 2009) also successfully applied EPR to model pipe blockages in Oslo (Norway). Their results indicated that blockage rate has a positive correlation with pipe age and strong negative correlations with pipe diameter, slope and total length.

2.4 Identified research gaps

Based on the reviewed literature, the primary identified research gaps are as follows:

- Appropriate methods for quantifying temporal and spatial variations of inflows to sewer systems are a prerequisite to effective sewer system modelling. Previously, different methodologies and models have been developed for generating flow and water quality determinants under dry weather flow (DWF) conditions. However, these methodologies pose some limitations if used for generating input time series for large scale modelling particularly in developing mega-cities where, due to cost and operational constraints, comprehensive field collection of data for characterising DWF outflows from large numbers of sewer sub-catchments is not feasible. Therefore, there is a research need for implementing and testing time series generators able to characterise DWF profiles from gauged and ungauged sub-catchments including uncertainty.

- Flow and water quality models of the urban drainage system can be useful to assess and manage its performance and to plan its development. However, due to data and computational costs, sophisticated, high resolution contemporary models of the sewer system may not be applicable. This constraint is, again, particularly marked in developing country mega-cities where catchments can be large, data tend to be scarce, and there are many unknowns, for example regarding sources, losses and wrong connections. Furthermore, there are many factors that affect the formulation of a suitable model for large sewer systems. For example, there are a variety of alternatives to represent system's spatial and temporal variability. Consequently, identifying parsimonious models for application to large and complex sewer systems for the purpose of providing inputs to WWTP operation is required.

- Due to increasing customer and political pressures, and more stringent environmental regulations, sediment and other blockage issues are now a high priority when assessing sewer system operational performance. Nevertheless, the complex physical mechanisms of sediment deposition, and the number of factors that may contribute to a sediment-related blockage, make forecasting a challenging task, and successful forecasting procedures are, in general, those based on statistical analysis of recorded sediment-related blockage events. Previous research in this area has been applied to relatively small urban areas and has identified the lack of studies at a much larger spatial scales. Thus complaint databases from large cities, containing historical records of system's performance, are considered to be especially valuable sources of data for further research.

2.5 Summary and conclusions

This concluding section presents how the previously presented literature review has defined the methods used in this Thesis. Key technical elements addressed within this dissertation can be summarised as follows:

- Despite the potential of stochastic models for synthesising WWTP influent both for flows and pollutant concentrations, this type of model has not been developed or tested extensively, particularly not for modelling DWF inputs to dynamic simulations of large sewerage systems with many ungauged areas. To evaluate the potential of this application, a methodology is presented in Chapter 4 which generates hourly diurnal variations of wastewater flow rates and pollutant concentrations. The models are generalised so that they can be used to predict DWF profiles from both gauged and ungauged areas. To represent the uncertainty in the daily pattern a stochastic model is used, allowing the uncertainty in characterising sub-catchment outflows to be modelled. Particular consideration is given to potential correlations between determinants and between sources because the combined effect of different sources and determinants is expected to affect the operation of the system. For example, the operation of a treatment works depends not only on the flow and BOD load individually, but also on the BOD concentration. This means that the determinants and the different sources should not be treated as independent and cross-correlation terms need to be included in the stochastic component of the model.

- There are many factors that affect the formulation of a suitable model for large sewer systems. Omitting fully distributed alternatives, general options are as follows: (a) a spatially averaged model without in-sewer processes representation, (b) a spatially averaged model with in-sewer processes representation, (c) a semi-distributed model with only advection-dispersion processes and (d) a semi-distributed model with advection-dispersion and in-sewer processes representation. To contribute to better understanding, Chapter 5 of this Thesis tests such options (a to d) with the aim of identifying a parsimonious model of dry weather pollutant dynamics for large and complex sewer systems for the purpose of providing inputs to WWTP operation. Besides this, it is still open the question of whether manual or automated sampling is suitable for characterising DWF dynamics within large sewer systems or if more advanced equipment, allowing a higher time resolution monitoring, are needed as well or instead. Consequently, an important part of this study is to review the value of the monitoring programme for supporting the modelling, and if possible make recommendations that may help direct future monitoring in Bogotá and similar studies.

- The very complex relationship between blockages and triggering mechanisms (Ashley et al. 2000) means that blockages often appear random with differences in blockage rates between catchments often inexplicable (Bachoc 1992; Laplace et al. 1992). Consequently, blockage forecasting is a challenging task and successful forecasting procedures are, in general, those based on statistical analysis of recorded sediment-related blockage events. Previous research in this area has been applied to relatively small urban areas and has identified the lack of studies at a much larger spatial scales. Thus complaint databases from large cities, containing historical records of system’s performance, are considered to be especially valuable sources of data for further research. Furthermore, this previous research has demonstrated that models using static variables are able to characterise pipe blockage rates. In principle, models can also include dynamic explanatory variables such as soil wetness and pipe structural condition to simulate the time-dependence of hydraulic deterioration (Tran et al. 2010). In general, however, it seems unlikely that sufficient dynamic data would be available to support this.

CHAPTER 3:
BOGOTÁ'S URBAN DRAINAGE SYSTEM:
Context, perspectives and relevant available data sets

As introduced in Chapter 1, one main case study is used for the development of this PhD dissertation: the urban drainage system of Bogotá's, the capital city of Colombia. Bogotá was proposed as case study with kind permission from the Bogotá Water Supply and Sewerage Company (Empresa de Acueducto y Alcantarillado de Bogotá - EAAB) (with gratitude to Mr. Juan Carlos Penagos) and the Environmental Engineering Research Centre (Centro de Investigaciones en Ingeniería Ambiental - CIIA) at the Universidad de los Andes (with gratitude to Professor Mario Díaz-Granados) to use their databases. Historically in Bogotá, efforts have focused on analyzing and improving the performance of individual components of the urban water cycle, without taking into account the interactions among them. Nevertheless, following the previously reviewed trends in developed countries (see Chapter 1), planning and operation is now starting its shift from the "*fragmented*" approach into an "*integrated*" one. Work at management institutions, environmental agencies, consultancy firms and different local universities, including monitoring, modelling, design and infrastructure development programmes, is currently conducted towards the implementation of an integrated management framework for Bogotá's urban drainage system. However, there are many important technical challenges to address in order to develop validated modelling tools that can be used within the decision making process. Some of these challenges have been identified and studied in the course of this research. In this Chapter, the main characteristics of the study area along with a detailed justification of the research and relevant available data sets are presented.

3.1 General description of the study area

Colombia (South America) has around 36 million inhabitants living in urban areas, equivalent to nearly 80% of the total Colombian population (WHO-UNICEF 2004). Bogotá, the capital city, is located on a fertile Andean plateau in the central region of the country at an approximate altitude of 2600 meters above sea level (m.a.s.l.) and surrounded by several municipalities (Figure 1). The Bogotá Savanna is the most densely populated area of the country, being an important industrial and agricultural region. Consequently, there are four main demands on water resources: water supply (domestic, commercial and industrial), hydropower electricity generation, natural drainage and agriculture. However, human consumption demands most of the available water. The population of Bogotá city increased by approximately 4 million in the last 30 years, from 2.9 million inhabitants in 1973 to 6.8 million in 2005 (DANE 2005). As part of this population growth, many families illegally settled in high risk areas, either on the foothills or along the creeks and rivers, facing lack of legal and effective water and sanitation services. It is estimated that the saturation population will be around 12 million inhabitants, accounting for about 20% of the total population of Colombia.

The Bogotá River basin has a total area of nearly 600,000 ha (CAR 2006). The Bogotá River, which starts at an altitude of 3300 m.a.s.l. from a pristine lake at the "Páramo de Guacheneque", drains the Bogotá Savanna along a course of about 330 kilometres until it joins the Magdalena River at an approximate altitude of 300 m.a.s.l. The monthly mean discharge of the Bogotá River varies from 1 m^3s^{-1} at the upper catchment to 40 m^3s^{-1} at the confluence with the Magdalena River, thus, it is a relatively small river receiving the wastewater discharge of about 8.5 million inhabitants. The river crosses eleven small municipalities before reaching Bogotá city. Most of these municipalities have a wastewater treatment system, commonly a facultative lagoon. However, Villapinzón, which is the first municipality in the upper catchment, has no wastewater treatment system despite producing highly contaminated domestic and industrial effluents from the leather industry. Additionally, the constructed treatment systems do not perform as designed. This is because of poor maintenance and operational capacity, and usually they do not treat all the wastewater generated due to deficiencies in the sewer and drainage systems.

While the Bogotá River flows from the Savanna to the Magdalena Valley, it crosses several other municipalities. Due to the heavily polluted condition of the Bogotá River (pathogens, organic matter, and toxic substances), the water provision for the municipalities located along its length is a problem of major concern. Despite the known problems with the Bogotá River water quality, some rural areas in these municipalities use the river as source of irrigation water. Until 2006, the Agua de Dios municipality in the downstream reach of the Bogotá River had used the river as water source for human consumption without a reliable treatment, imposing health risks to a large population. Since 2006, this municipality has used water from the Magdalena River. It is clear that there is a gap between the desired uses of the water along the river and the quality required to support them (Uniandes 2005).

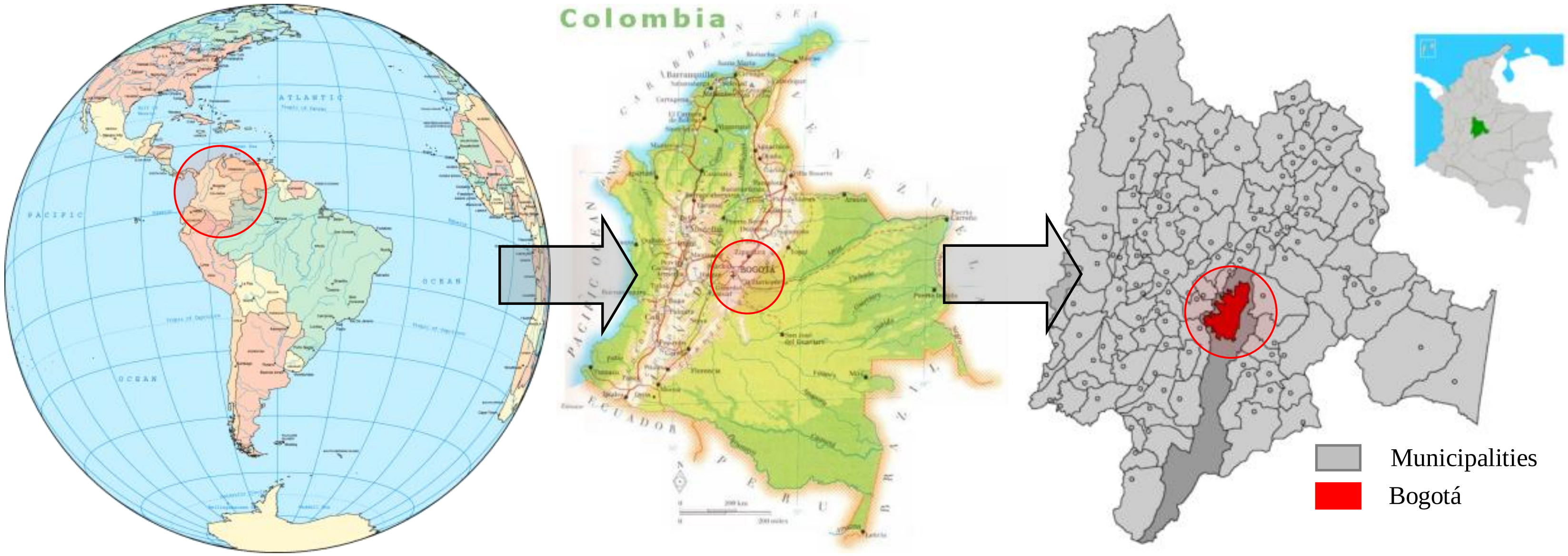


Figure 1. Location of Bogotá

3.2 Bogotá's urban drainage system

3.2.1 The sewer system

The Bogotá sewer system consists of three main catchments: Salitre, Fucha and Tunjuelo, together covering approximately 330 km^2 (Figure 2). Each sub-catchment is drained from east to west by one urban river with the same name as the catchment. There are also three relatively minor catchments named Torca, Tintal and Soacha. The drainage system is a combination of the two traditional types of sewer system - combined and separated. The first sewer system developments, in the central part of Salitre sub-catchment and the eastern part of Fucha sub-catchment, consist of a combined system draining approximately 74 km^2, while the newer systems (since 1965) are separated. Approximately 1.3 million inhabitants are served by combined systems (nearly 20% of total population). The main problems related with the combined system in Bogotá can be summarized as follows: direct discharges of wastewater without treatment as a consequence of the absence or lack of adequate infrastructure, and CSO discharges even during dry weather periods. Regarding the separated system, there are many wrong connections from the wastewater system flowing into the storm drainage system. Water quality measurements in the open channels for storm drainage in the Fucha sub-catchment indicated that up to 22% of connections from the wastewater system entered the channels (Grucon-IEH-Soprin 1999). There are also wrong connections from the storm drainage system flowing into the wastewater system. Based on flow surveys in some interceptors, from 23% up to 90% of connections from the storm drainage system enter the wastewater system (Grucon-IEH-Soprin 1999). As a consequence, water quality in the storm drainage channels is bad and there is a high risk of sewer flooding during intense rainfall events. Actually, the separate system acts more as a "dual" combined system rather than as a separate one.

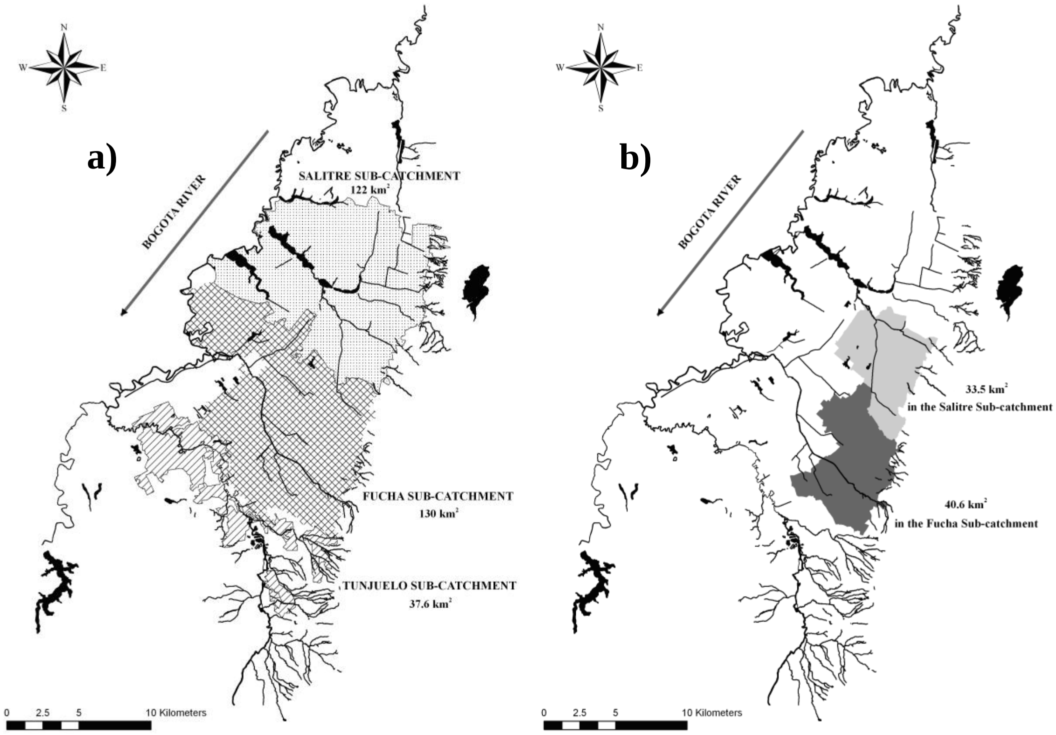


Figure 2. Bogotá's urban drainage system: (a) Main sewer system sub-catchments; (b) Combined sewer system in Salitre and Fucha sub-catchments

3.2.2 The wastewater treatment plant

In many developed countries the wastewater infrastructure is nearly complete, in the sense that almost all the sewage is collected and treated to a certain extent. In contrast, Bogotá city only has one WWTP in operation. In the last three decades, several sanitation alternatives have been proposed for Bogotá city. By 1994 the local administration decided to adopt a strategy based on a series of treatment plants. The implementation of the sanitation scheme started without the consultation of the city council, national government and general public, and with financial restrictions which would not allow the full proposal to be realised. It was facilitated by a provision of an environmental law approved in 1993 which assigned additional economic resources to the urban environmental authorities. The fact that the urban environmental control agency (named Secretaría Distrital de Ambiente - SDA) had the responsibility of implementing the selected sanitation scheme generated conflicts of interests. Additionally, the decision to proceed with the plan was not the result of an appropriate evaluation and it ignored the results of previous technical studies. The decision was not coordinated with the sanitation program for the upper catchment of the Bogotá River (including 23 wastewater treatment plants) which had been implemented by the Regional Autonomous Environmental Agency of Cundinamarca (CAR). In 2000, the first stage of one of the proposed treatment plants (the Salitre WWTP) entered into operation under a Build-Operate-Transfer concession contract with a private international consortium. In late 2003 the local administration cancelled the contract, purchased the plant from the consortium, and turned over its operation to the EAAB. As no significant improvement to the Bogotá River water quality was noticed, the sanitation scheme was modified. In 2007 the national environmental authority issued an environmental licence to EAAB to expand the Salitre WWTP. This expansion project is arranged to be financed by the CAR, as it is planned that the Salitre WWTP will supply water to the 6,000 ha "La Ramada" irrigation district which is under CAR administration.

The Salitre WWTP was built between 1997 and 2000. It has 4 m^3s^{-1} of treatment capacity (roughly 25% of the total amount of wastewater generated in Bogotá) using physical/chemical treatment. Local environmental standards demand a treatment performance of at least 40% for the organic load and 60% for the sediment load. The developments agreed in 2007 involved plans to upgrade the Salitre WWTP (increasing the capacity to 8 m^3s^{-1} and adding a secondary treatment) and build a new secondary WWTP to treat the wastewater produced in the Fucha and Tunjuelo sub-catchments. For the operations of these new installations, a good understanding of the flow and water quality characteristics of the Bogotá urban drainage system is required.

3.2.3 The receiving water courses

Direct discharge of urban dry weather wastewater flows into receiving waters is no longer generally acceptable (Vollertsen and Hvitved-Jacobsen 2000). However, this type of pollution source is very common in the Bogotá River catchment watercourses due to the absence of appropriate water treatment infrastructure. The fact that the self-purification capacity of the Bogotá River is very limited in the middle catchment (the area of influence of Bogotá city) due to low flow (low dilution capacity), the small longitudinal slope (0.006%, limiting reaeration processes), high altitude (2556 m.a.s.l.) and medium temperature (decreasing the saturated dissolved oxygen concentration) worsen the situation (Camacho et al. 2002).

Models of the water quality in the Bogotá River offer the opportunity to simulate the effects of improvements to the urban wastewater drainage and treatment system, hence are valuable planning tools for evaluating and optimising different strategies. Since early 2000s, great efforts have been carried out in order to implement water quality models in the receiving water courses: the Bogotá River and the Salitre, Fucha and Tunjuelo Rivers. For example, during a two year research project (which started in 2001), five field campaigns were carried out, covering different hydrological conditions for the Bogotá river (Camacho et al. 2002). The results, including a stretch of 60 km under anaerobic conditions and BOD concentrations of up to 140 mg l^{-1}, illustrate the main problem of the Bogotá River. The (steady state) modelling results obtained by Camacho et al. (2002) clearly demonstrate that a sanitation plan involving the construction of three secondary treatment wastewater plants[1] was not sufficient to recover the water quality of the Bogotá River and to solve the public health problems caused by bacteriological contamination (mean Total Coliform concentration is around $1x10^8$ MPN/100 ml). This is because Bogotá city is not the only source of pollutants. In the upstream reach, the Bogotá River receives direct wastewater discharges from other urban settlements and diffuse loads from agricultural activities. Due to natural reaeration provided by "Tequendama's Fall" (the altitude decreases by approximately 1900 m in a very short river stretch), the DO concentration after the fall is near saturation, around 6 and 7 mg l^{-1}. Water from the Bogotá River is pumped into the Muña Reservoir at the end of the middle catchment, and it is continuously used for hydropower generation by the Pagua's chain with a capacity of 600 MW, taking advantage of the fall that exists between the Bogotá Savannah and the low part of the river basin. The water from the old Casalaco's hydropower

[1] One WWTP per sub-catchment (Salitre, Fucha and Tunjuelo)

generation chain with a total capacity of 95.6 MW, which currently is operated a few times during the year, continues also by the waterway of the Bogotá River. Both processes of hydropower generation contribute to the increase of DO in the Bogotá River due to the high slopes, the falls, and the action of the Pelton turbines (Uniandes 2005). DO in the down-stream reach of the Bogotá River then decreases by up to 0.5 mg l^{-1} as a consequence of nitrification and organic matter degradation processes (Camacho et al. 2002). Currently, the water quality of the Muña Reservoir is a main concern. The reservoir receives high pollution loads (organic matter, nutrients, pathogens and toxic substances) from the Bogotá River, and an excessive growth of floating aquatic plants appears. These aquatic plants tend to cover the water surface, and generate problems of odours, mosquitoes and corrosion of the electrical generation equipment. However, the Muña Reservoir generates a net positive impact on the water quality of the lower basin of the Bogotá River due to sedimentation of organic matter and suspended solids, and therefore of pathogenic organisms and toxic substances associated with the solids (Uniandes 2005).

3.3 Contribution of this Thesis to the case study

For a better understanding about how this PhD research can contribute to the challenging Bogotá case, this section aims at describing the current local water management plans. In general, Colombia has made significant progress towards improving urban water and sanitation coverage (WHO-UNICEF 2010; WHO-UNICEF 2010). However, wastewater treatment rate is still very low, with only about 22% of municipalities having partial treatment facilities and it is estimated that only 10% of them perform adequately. In most of large urban centres, only 20% to 25% of the produced wastewater is treated to some extent (MAVDT 2004). Due to the urgent need to implement wastewater treatment schemes across the country, in 2004 the Ministry of Environment, Housing and Territorial Development (MAVDT) formulated a framework to rapidly develop wastewater programs (MAVDT 2004). Particularly for the Bogotá river basin, the National Planning Department (DNP) issued a strategic planning document which called for upgrading and extending the Salitre WWTP, the construction of a large plant downstream of Bogotá, the expansion of wastewater treatment installations in the upper basin and the mitigation of environmental impacts related with the Muña Reservoir (DNP 2004). In 2006, the CAR released Bogotá's River basin plan (CAR 2006) which established the water quality standards and targets (for year 2020) for the complete basin. The objective is to achieve raw water quality standards suitable for agricultural use in Bogotá's River middle basin. To complement the basin plan, each municipality has to agree a plan for sanitation and wastewater discharge management in accordance with the water quality targets for the Bogotá River. For Bogotá city, the SDA approved EAAB's plan for sanitation and wastewater discharge management (SDA 2007). The plan for sanitation and wastewater discharge management, which is part of the most recent water supply and sewer system master plan (EAAB 2006), primarily establishes targets (from 2007 until 2017) for pollutant load reduction (i.e. BOD and TSS loads) in order to improve the water quality status of Bogotá's urban rivers. As part of this, it is also requested actions regarding urban drainage proactive maintenance and system rehabilitation.

Recently, three comprehensive hydraulic, solute transport and water quality measuring campaigns along the whole main branch of the Bogotá River were carried out by the Universidad Nacional de Colombia (Bogotá) and EAAB which was the basis for sensitivity analysis, calibration and validation of a detailed dynamic Bogotá River water quality model (UNAL 2010). This dynamic model was further used in a recent research study which assessed the expected water quality status of the Bogotá River when: (*a*) the extended 8 m^3s^{-1} Salitre WWTP and the additional future WWTP enter into operation and (*b*) ongoing sewer developments (i.e. construction of main sewer collectors to transport wastewater to the new WWTP), part of the plan for sanitation and wastewater discharge management, are functional. It was concluded that it is necessary to implement more advanced treatment at both WWTPs, including disinfection and nutrient removal. It was also highlighted the urgent need of the implementation of sustainable urban drainage strategies such us source control in order to achieve the Bogotá River water quality targets (González 2011).

With that context as background for this Thesis, repeating the aims listed in Chapter 1: (*a*) Chapter 4 aims at an improved understanding of pollutant load contribution at the sub-catchment level under dry weather conditions, thus providing tools to prioritise potential source control strategies; (*b*) Chapter 5 presents a validated model to assess and improve the interaction of the design and operation of the sewer system and the Salitre WWTP under dry weather conditions and (*c*) Chapter 6, proposes and critically assesses a novel statistical method supported by an exceptionally long and spatially detailed customer complaints data-base to support proactive maintenance of the sewer system. Results presented in this dissertation not only aim to improve the technical knowledge for the particular context of Bogotá but to provide advances in science that can be useful for other, both developing and developed, countries.

3.4 Description of available data sets

To support model and method development, different databases have been made available thanks to existing research cooperation between the author of this dissertation, the EAAB and the Universidad de los Andes as mentioned above. Some of these databases have been acquired under active participation of this dissertation author before and during the PhD period (particularly the data used in Chapters 4 and 5). The specific activities (if any) in which this dissertation author was directly involved are listed along the description of the data sets.

3.4.1 Available data sets to characterise dry weather loads to sewers

In recognition of the importance of observations for supporting urban drainage planning and operations, a major data collection programme has taken place in Bogotá since 2006 by the Environmental Engineering Research Centre at the Universidad de los Andes with support from EAAB (Díaz-Granados et al. 2009). A database which now covers quantity and quality from about 150 monitoring sites in Bogotá's sewer system for primarily dry weather conditions is available. Field data were collected in four monitoring periods: (*i*) 12 sites in 2006, (*ii*) 17 sites in 2007, (*iii*) 82 sites in 2009 and (*iv*) 36 sites in 2011. For all the monitoring periods, Juan Pablo Rodríguez was involved in the design of the monitoring works (e.g. definition of sites, sampling methods, duration of activities). Particularly for periods (*i*) and (*ii*): he led the group of field technicians that carried out the monitoring campaigns, interacted with the laboratory technicians that analysed the wastewater samples, validated the laboratory reports and helped in the design and implementation of a structured database which include in-situ and laboratory data.

The main focus of the sewer system monitoring programme was on flow, BOD, COD and TSS as local environmental regulation focused on them when assessing wastewater discharges and WWTP performance. All wastewater samples were characterised in the EAAB laboratory following the "Standard Methods for the Examination of Water and Wastewater", the SM-5210 for BOD, SM-5220 for COD, SM-2540 for TSS and SM-4500 for DO (see Standard Methods Committee).

There are 106 combined sewer and 282 separate sewer sub-catchments in Bogotá. Currently, observed in-sewer flow and quality data from about 150 sites are available as described above. However not all sites characterise single sub-catchments and hence a suitable sub-sample of the data was selected for the modelling of sub-catchment outflows. 24-hour period time series from 29 sites (Figure 3) are used. The rest of the data correspond to downstream sites and are being used for calibration and validation of the larger sewer system model (see next sub-section). Monitoring activities were carried out one site at a time during dry weather week days over a single 24-hour period per site. Three sites (out of the 29) were monitored over one additional 24 hour period. A dry weather day is defined here by the following criteria: (*a*) it corresponds to a non-holiday weekday; (*b*) both quantity and quality measurements are available, (*c*) it lasts from 00:00 to 24:00, (*d*) no precipitation has been observed during the considered day and (*e*) there is no evidence that rainfall events have influenced neither the initial nor the final conditions of flows and water quality. Monitoring characteristics are presented in Table 2.

Table 2. Details of dry weather wastewater quality collected and used in this study

Phase Number	Sampling resolution	Measurement sites	Samples per site per determinant (over a period of 24 hours)
2006	Hourly	2	24
2007	Variable	4	12
2009	Hourly (composite)	13	8
2011	Hourly (composite)	10	8

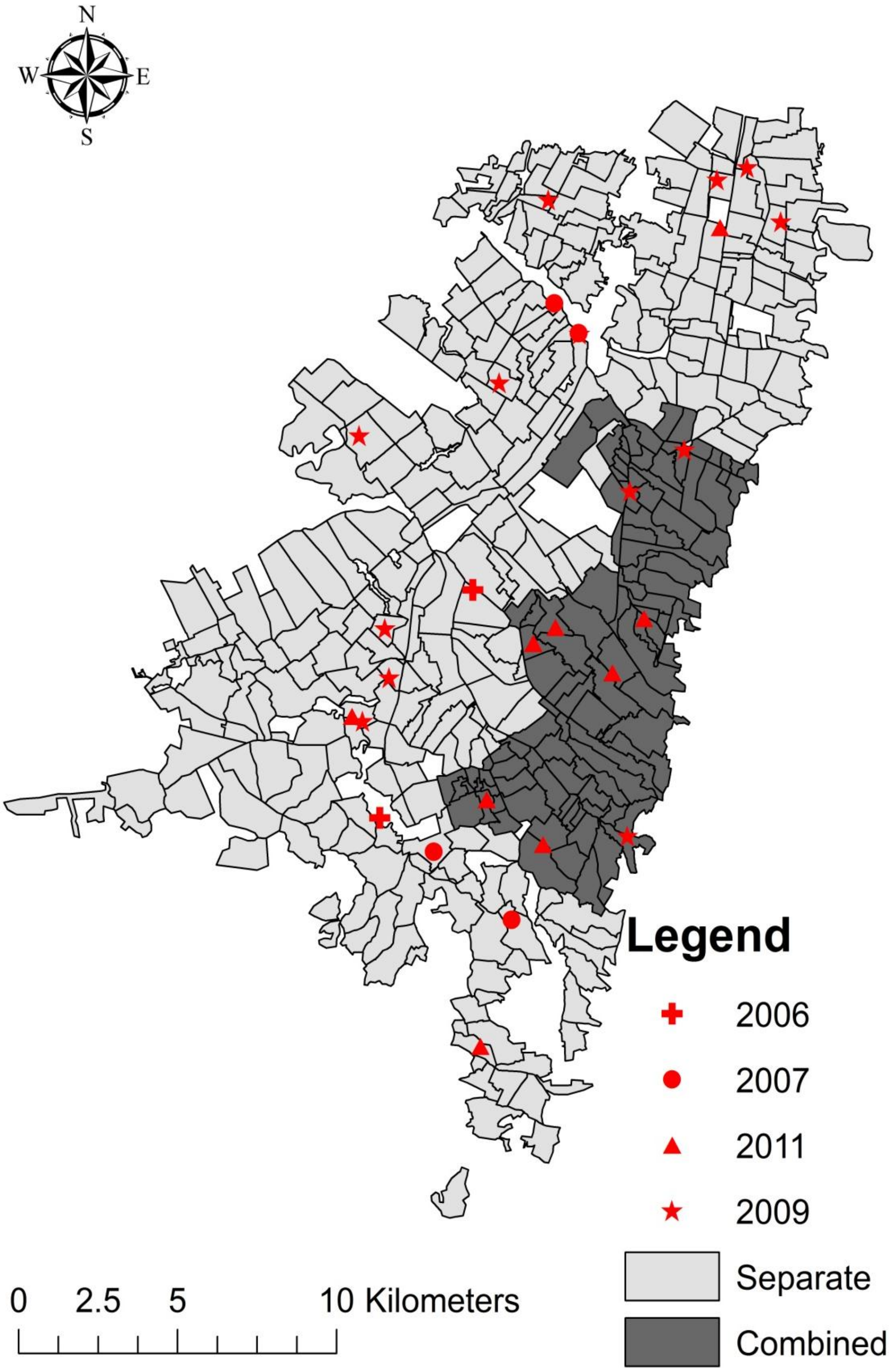


Figure 3. Monitoring sites used in 2006, 2007, 2009 and 2011 field campaigns

For the Bogotá case a detailed consumption database allows the estimation of the number of residential, industrial, commercial, office-use and multi-user (i.e. apartment blocks) units in each sub-catchment; and the average daily water consumption per unit for each of these five components. Each unit is an individual property which has a water supply connection and contract. The water use data for 2006 were used and considered to be representative of the field campaign conditions. Based on all sewer sub-catchments in Bogotá, Figure 4 shows the probability density functions for area and inhabitants and the cumulative distribution functions for water use disaggregated into residential, industrial, commercial, office-use and multi-user components, as percentages of total water use. It can be seen that the average sub-catchment has an area of about 85 ha, approximately 18,000 inhabitants and water use is predominantly residential. Even though most of the sub-catchments are mainly residential, there are areas where industrial consumption reaches 70 % of total consumption. The 29 gauged sites used in Chapter 4 and Chapter 5 have the following characteristics: catchment areas range between 6 and 197 ha with a mean of 92 ha; and mean daily flows between 0.002 and 0.081 m^3/s with a mean of 0.03 m^3/s. Regarding the variation in water use between sub-catchments, residential use varies between 3 % and 100 % with a mean value of 77 %; industrial use between 0 % and 70 % with a mean value of 9 %; commercial use between 0 % and 61 % with a mean value of 11 %; and office-use between 0 % and 14 % with a mean value of 3 %. Figure 5 summarises the number of water consumption units per hectare for each of the 29 monitored sub-catchments, Figure 6 presents the corresponding box-plots for observed sub-daily BOD, COD and TSS concentrations and Figure 7 presents box-plots of observed daily mean BOD, COD and TSS concentrations at the 29 gauged sub-catchment outlets. In Figure 7 is evident the high spatial variability of observed mean values. Over the 29 mean values per determinant, mean values and related standard deviations (in brackets) for BOD, COD and TSS are respectively: 296 (123), 619 (241) and 193 (85), all values in mg l^{-1}.

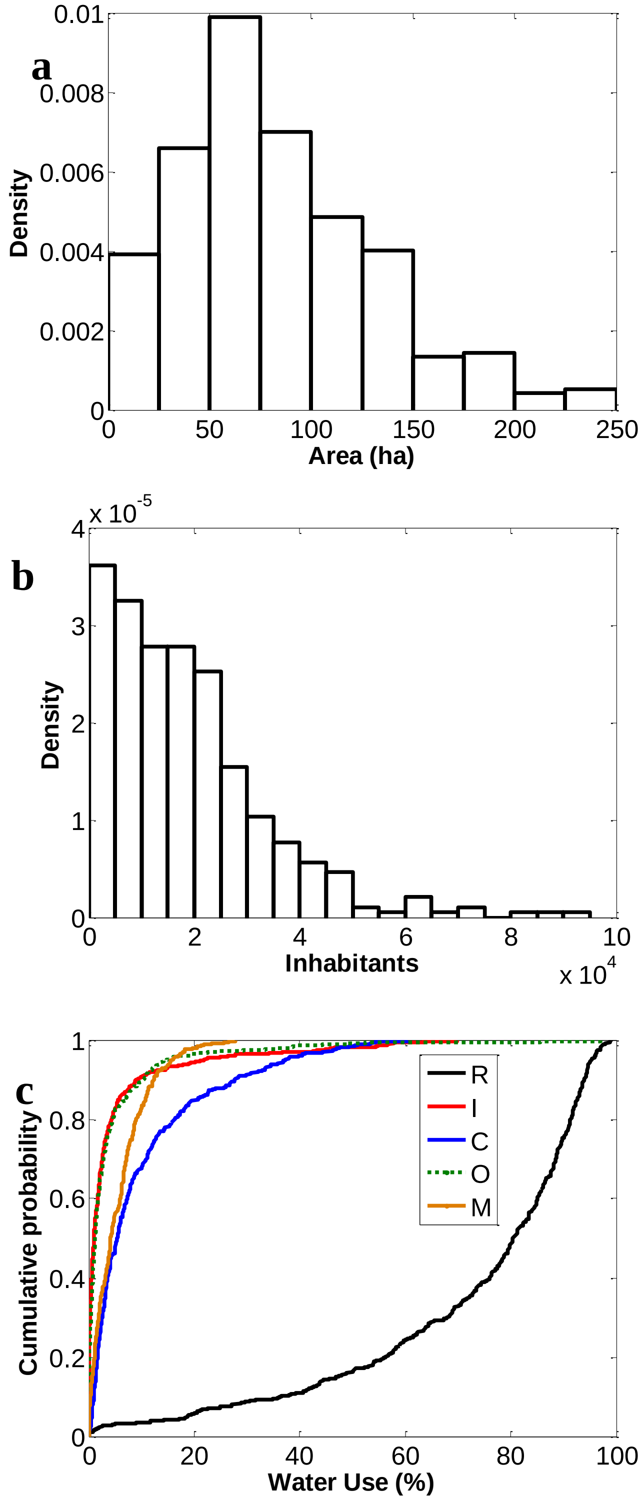


Figure 4. Characteristics of Bogotá's urban sub-catchments: (a) probability density function for sub-catchment area (in ha), (b) probability density function for number of inhabitants; and (c) cumulative distribution functions for water use type, where R = Residential, I = Industrial, C = Commercial, O = Office and M = Multi-user

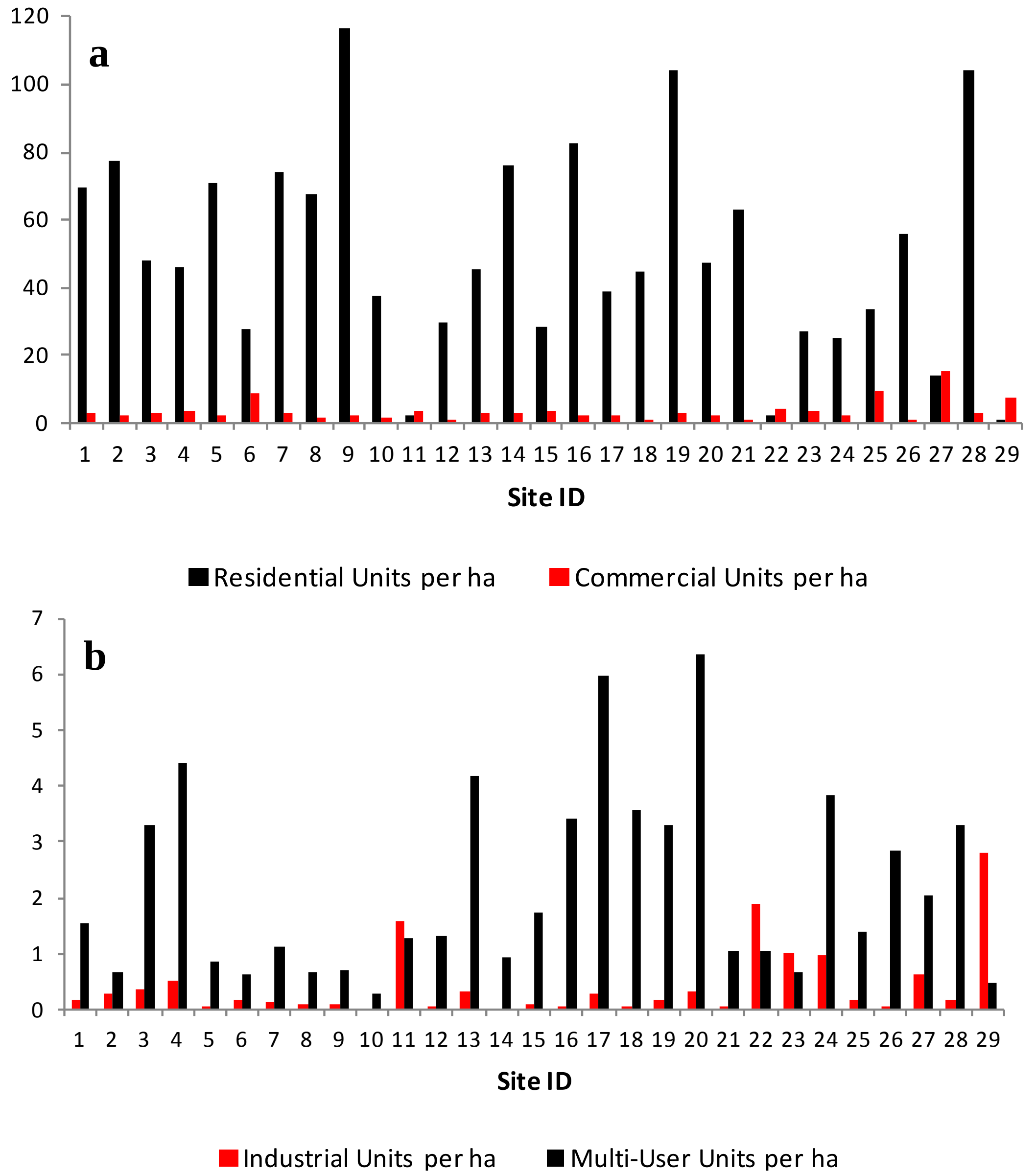
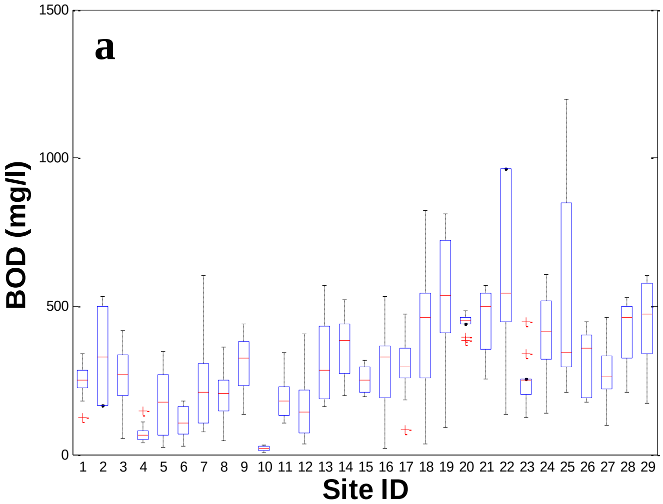
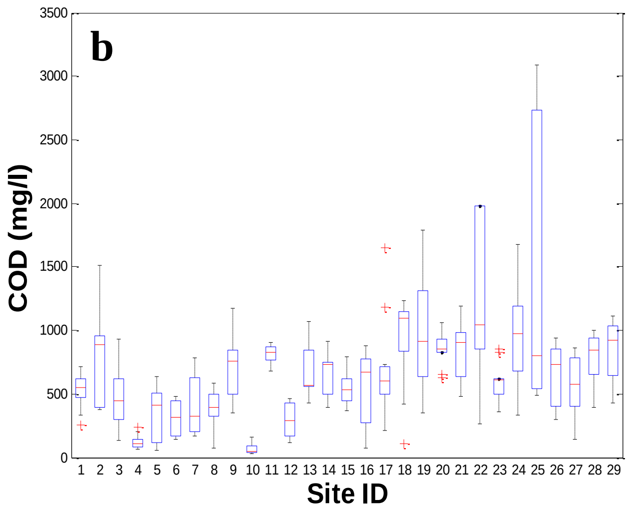


Figure 5. Number of water consumption units per hectare for each of the 29 monitored sub-catchments: (a) residential and commercial units and (b) industrial and multi-user units





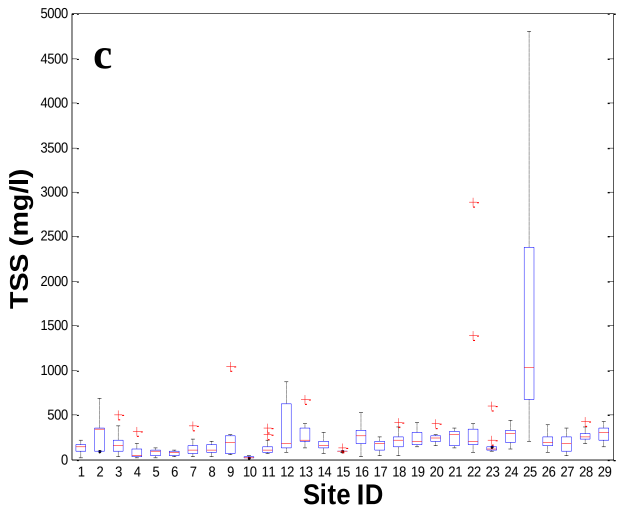


Figure 6. Measured (a) BOD, (b) COD, (c) TSS concentrations (mg/l) box-plots at 29 single sewer sub-catchments in Bogotá

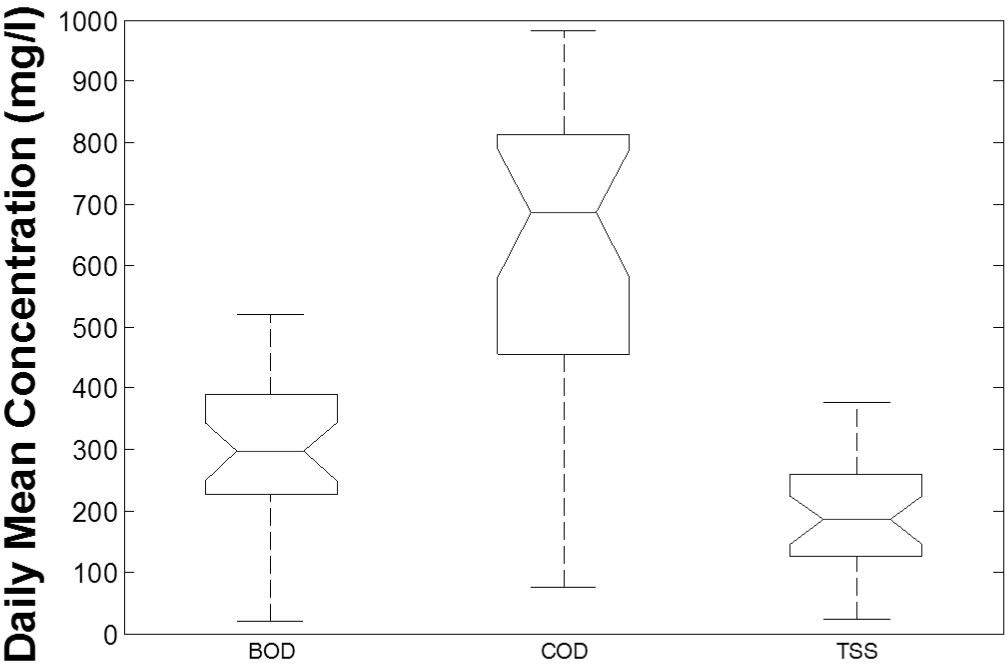


Figure 7. Daily mean concentrations of BOD, COD and TSS (estimated over 29 sub-catchments in Bogotá)

3.4.2 Available data sets to support large scale sewer system modelling

The focus of Chapter 5 is on the sewer system connected to the Salitre WWTP (see Figure 8), which corresponds to a 150 km^2 urban catchment (also named Salitre) serving nearly 2.5 million inhabitants (about 1/3 of Bogotá's population). The Salitre catchment is composed of around 180 sub-catchments. The catchment is predominantly residential (residential water consumption corresponds to 84% of the total, with 16% standard deviation in this value between the 180 sub-catchments) with moderate industrial and commercial activities (on average 3% and 11% of the water consumption respectively). Nevertheless, industrial and commercial uses in particular areas reach up to 45% and 54% of the sub-catchment water consumption respectively. Figure 9 shows some statistics to further introduce the case study: (*a*) the probability density functions for sub-catchment area (in ha), inhabitants and mean slope (m/m) and (*b*) the cumulative distribution functions for sub-catchment water use disaggregated into residential, industrial, commercial and office-use components as percentages of total water use.

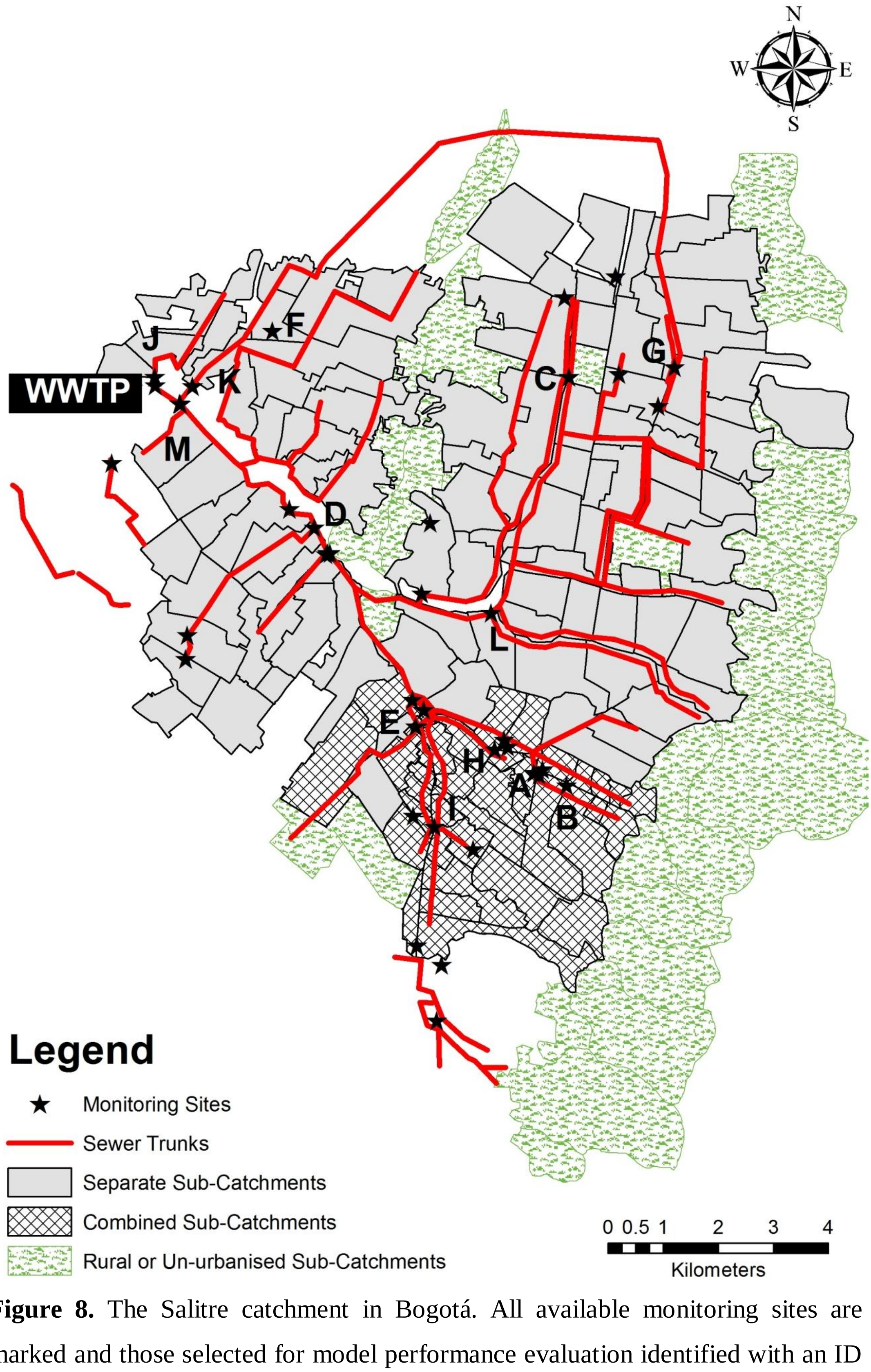


Figure 8. The Salitre catchment in Bogotá. All available monitoring sites are marked and those selected for model performance evaluation identified with an ID letter

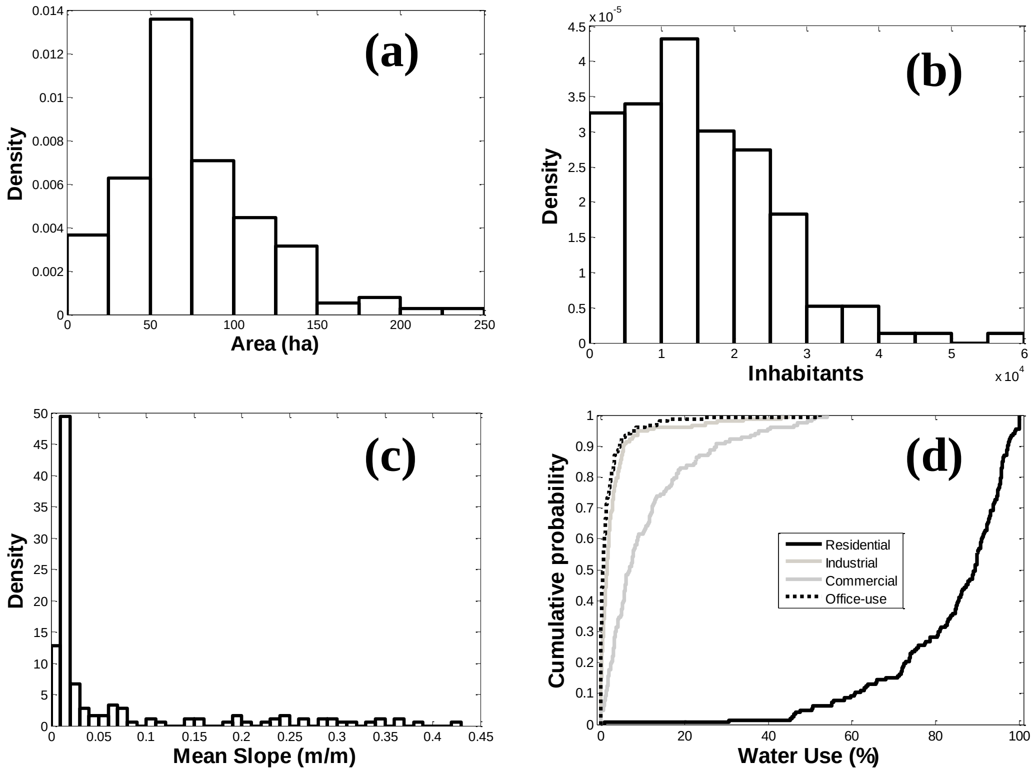


Figure 9. Properties of the Salitre sub-catchments: (a) Probability density function of areas; (b) Probability density function of inhabitants; (c) Probability density function of mean surface slope (m/m); and (d) Cumulative distribution function for water use disaggregated into residential, industrial, commercial and office-use components

Data from gauged sub-catchments and downstream locations in the Salitre catchment are used for calibration and validation of the model (see details in previous sub-section). 13 representative gauged sites were selected for assessment and validation of the model and identified with ID letters in Figure 8. They include combined and separate (both wastewater and stormwater systems) sub-catchment outlets, CSOs and sites within the main sewer trunk system (see Table 3). The 13 sites allow quantification of the model's ability to represent flow and quality at structures such as CSOs and, to some extent, quantify the flow balance effects of wrong connections. These sites also correspond to different maximum hydraulic retention time, varying from 0.6 up to 6.7 hours (these retention times account, if applicable, for both transport within sewer sub-catchments and main sewer trunks). Particularly for sites representing sub-catchment outlets, the main characteristics are: site A has 41 ha, about 6,000 inhabitants and commercial water use reaches 13% of total consumption; site C has 59 ha, about 30,000 inhabitants and industrial and commercial water uses reaches 4% and 8% of total consumption; site D has 109 ha, about 50,000 inhabitants and commercial water uses reaches 11% of total consumption; site F has 82 ha, about 20,000 inhabitants and residential water uses is 96% of total consumption; site G has 73 ha, about 20,000 inhabitants and residential water uses is 96% of total consumption.

Table 3. Characteristics of 13 monitoring sites selected for model validation

Monitoring Site	Description	Hydraulic Retention Time (hours)
A	Combined sub-catchment outlet	0.6
B	Main sewer trunk	0.7
C	Separate sub-catchment outlet (Wastewater System)	0.7
D	Separate sub-catchment outlet (Stormwater System)	0.7
E	Main sewer trunk	0.8
F	Separate sub-catchment outlet (Wastewater System)	1.0
G	Separate sub-catchment outlet (Wastewater System)	1.0
H	Combined sewer overflow (CSO)	1.1
I	Combined sewer overflow (CSO)	1.3
J	Main sewer trunk (outlet flowing into WWTP)	2.3
K	Main sewer trunk (outlet flowing into WWTP)	4.1
L	Main sewer trunk	5.6
M	Main sewer trunk (outlet flowing into WWTP)	6.7

As reviewed in Chapter 2, DO and temperature play a fundamental role in wastewater transformation processes. Besides this, the Salitre catchment has predominantly mild slopes (see Figure 9c). Consequently, in-sewer system reaeration is not expected to be high thus DO levels are also expected to be low. For a better understanding of expected DO and temperature within the El Salitre sewer system, and therefore to identify to what extent it is beneficial to explicitly consider complex transformation processes, additional available measurements for the Bogotá case are described. From monitoring campaigns carried out by the Universidad de los Andes and EAAB data from 20 untreated wastewater discharges to water courses (ranging from single sub-catchment outlets to large sewer trunks) are available. Juan Pablo Rodríguez was in charge of the monitoring campaigns that provided DO and temperature data. He was involved in the design of the monitoring works, led the group of field technicians that carried out the monitoring campaigns, interacted with the laboratory technicians that analysed the wastewater and water samples, validated the laboratory reports and helped in the design and implementation of a structured database which include in-situ and laboratory data. These campaigns were part of a detailed characterisation of the receiving water courses which drains the Bogotá city. Different untreated wastewater discharges into the rivers and stations along the watercourses were monitored.

The previous dataset is not used for model evaluation in Chapter 5 as it corresponds to one instantaneous sample thus not providing enough information about system dynamics. In total 40 samples (grabbed at different times through week days, ranging from 7:00 to 17:00) were characterised using the corresponding standard method with a detection limit of 0.1 mg l^{-1}. It was found that for single sub-catchment discharges (15 sites) the mean DO concentration was 1 mg l^{-1} (standard deviation 1 mg l^{-1}) while for all main sewer trunks discharges (5 sites) DO was not detectable in any case implying concentrations lower than 0.1 mg l^{-1}. Across all measurements, temperature varied between 17 °C to 30 °C (with mean and standard deviation of 20 °C and 3 °C respectively). These results imply that the reaeration process is unlikely to be a main driver of in-sewer transformations and that due to relatively high temperatures anaerobic conditions are expected to prevail within the system.

Apart from previously described databases, the EAAB collected wastewater samples providing simultaneous and continuous 4 hour time resolution flow, BOD, COD and TSS at different locations in the Salitre catchment. Campaigns carried out in 2009 (2 August to 15 September) and 2010 (4 to 24 February) correspond primarily to DWF conditions and are used in Chapter 5. Some of the locations coincide with sites reported in Table 3 and presented in Figure 8. In this work, 3 of these locations were selected: (a) a site downstream E (with a maximum retention time of about 1.8 hours), (b) a site downstream L (with nearly the same maximum retention time as site L – 5.6 hours – but collecting additional wastewater contributions) and (c) site M. Only data of week days (23, 20 and 13 days at (a), (b) and (c) respectively) are used for analysis.

3.4.3 Data sets to support proactive management of sediment-related sewer blockages

Bogotá's sewer system is divided into five operational zones for the purpose of sewerage management. For the sake of the research study presented in Chapter 6, these zones are further split into a total of 24,392 grid squares of about 0.03 km^2 each (Figure 10). For each square, the number of manholes (191,662 in total), number of gully pots (136,252 in total) and total length (totalling 7,678 km), mean diameter and mean slope of pipes (in both stormwater, foul and combined systems) are recorded. There are more uncertainties in the slope data compared to the rest of the properties-related data: nearly 12% of sanitary and 16% of stormwater pipes have unknown slopes. Nevertheless, there is still a considerable fraction of pipes (more than 85%) with known slopes.

The record of sewer failures for each square is derived from a platform for collecting and handling customer complaints run by Bogotá's water supply and wastewater utility. A dedicated centre receives customer calls informing about sewer system failures. Before a reported failure is verified, the following information is recorded: event identification code, address, geographical coordinates, report date and time, unconfirmed failure type, and type of hydraulic structure in which the failure has been observed (manhole, gully pot, pipe, etc.). After field verification, the reported failure is classified as 'effective' or 'ineffective'. If effective, a more detailed identification code for the failure type is reported, including the required corrective action. Once corrected, the finishing date and duration of the triggered corrective activity are added to the database. As historical records do not contain information about which particular hydraulic structure has failed, and if they did so it is unlikely that enough failures would be recorded to perform a useful statistical analysis, the aim of the modelling framework described in Chapter 6 is to understand failures within each square and within each hydraulic structure type. The customer complaints database for a 7.5 year period (January 2004 to June 2011) was available for the case study, which is the first published use of this data set in its full spatial scale.

An area of nearly 290 km^2, comprising 9,658 squares (out of 24,392) contains all sediment-related reports, which are 248,631 in total. These include ineffective (23%), repeated (17%), date-lacking (0.8%) and wrongly classified failures (4.2%). In order to achieve this level of maintenance, Bogotá's water utility had 16 Vactor 2100 sewer cleaners, 21 vehicles equipped with either rods or winchers and 14 dump trucks (values for 2007). During an average week day, each Vactor cleaner and dump truck was operated by 3 technicians, attended 7.6 verified reported sites, travelled 38.43 km and cleaned 250 m of pipes. It was found that 54%, 24% and 22% of effective blockages were related to pipes, gully pots and manholes respectively. Figure 11 shows some statistics to help further introduce the problem. Figure 11a presents the cumulative probability of the number of structural units (i.e. number of manholes, number gully pots and the total length of pipes/100 m) in those squares which have at least one effective blockage in the 7.5 year period. Apparently there are squares with no units of a particular type, yet there are blockages reported for that type. For example, Figure 11a shows that about 3% of the squares with gully-pot sediment blockages have no records of gully pots in the square.

This preliminary analysis of blockage data helped in identifying squares where either the official database of sewer structures or the geographical coordinates of the reported blockages should be improved. Figure 11b shows the cumulative probability of the number of effective blockages per square for each type of structural unit (i.e. manholes, gully pots, pipes and grouping all of them). Figure 11c reports cumulative probability values for the number of blockages per square per year for each structural unit. Previous literature includes no data on sediment-related blockages rates in manholes - blockage rates are generally reported solely for pipes and rarely for gully pots (ten Veldhuis and Clemens 2011). Figure 11 also illustrates the magnitude of the sediment problem in Bogotá: Figure 11c shows that over all squares the median value in Bogotá is about 1.5 blockages per km of pipe per year (the pipe structural unit is 100 m in Figure 11), which is comparable with the UK average pipe blockage rate that ranges between 0.1-2 blockages per km per year (Arthur et al. 2008).

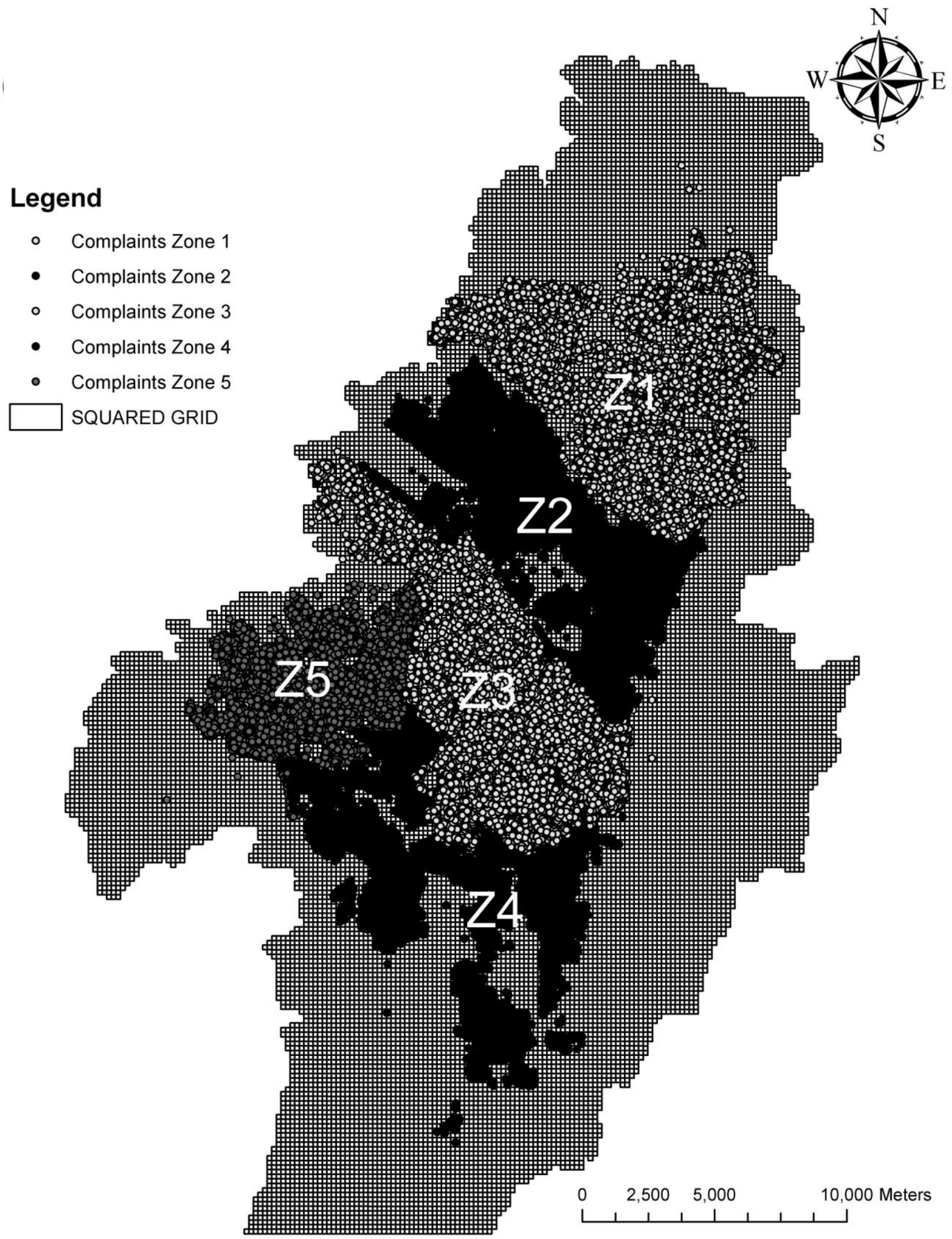


Figure 10. Implemented squared grid and spatial distribution of customer complaints for all five operational zones in Bogotá

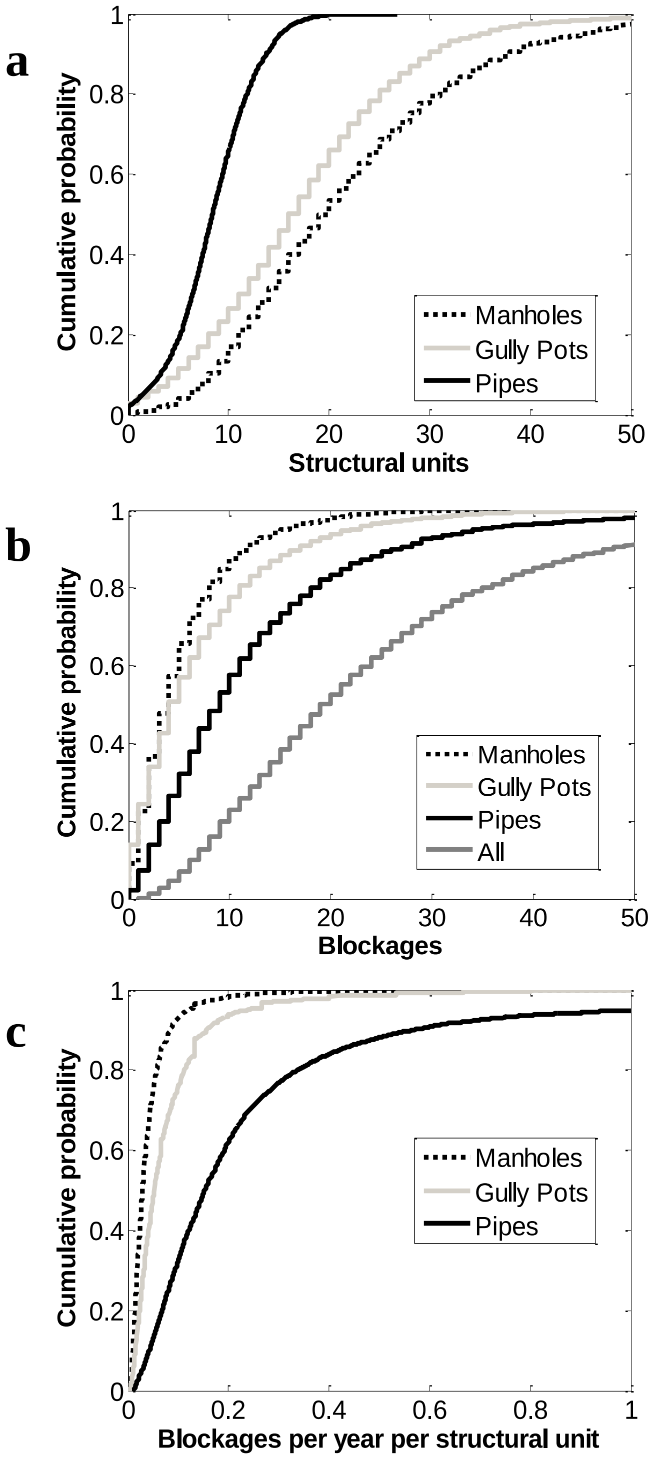


Figure 11. Available blockage data characterisation and preliminary analysis (all at square level): (a) cumulative probability of the number of structural units; (b) cumulative probability of the number of effective blockages for each type of structural unit and (c) cumulative probability values for the number of effective blockages per year for each type of structural unit. The pipe structural unit is 100 m length of pipe

3.5 Summary and conclusions

From previous sections in this Thesis, it can be concluded that Bogotá city provides a particular case study for urban drainage modelling: (*a*) it corresponds to a developing country urban centre experiencing relatively fast urbanisation processes (annual growth rate of 1.4%) thus having severe water quality problems in the watercourses due to lack of wastewater treatment facilities infrastructure, and a traditional separated wastewater management approach, (*b*) urban drainage planning and operation is now (conceptually) starting its shift from the "fragmented" approach into an "integrated" one: as part of this, the urban drainage system components (i.e. the sewer system, treatment facility and receiving water courses) has been (relatively) intensively monitored and (*c*) a major investment in the urban drainage system is expected in the near to medium term.

Despite of the advantages of some data availability and a local consensus regarding the need for an integrated urban drainage management approach, there are many important technical challenges to address in order to develop validated modelling tools that can be used within the decision making process. Some particular sewer system-related issues have been studied in the course of this research (as detailed in Chapter 1 and Chapter 2). This PhD dissertation can be seen as complementary research to other recent monitoring and modelling programmes carried out by different research groups at leading local universities (i.e. Universidad de los Andes, Universidad Nacional de Colombia – Sede Bogotá and Pontificia Universidad Javeriana). Such programmes are greatly contributing to a better understanding of all the integrated system components: the sewer system (Duarte et al. 2009; Estupiñan et al. 2009; Torres et al. 2011), the receiving water courses (Uniandes 2003; Raciny et al. 2008; UNAL 2010) and the treatment facilities (Rada and Torres 2011). Besides this, other recent research efforts looked at the assessment of holistically-oriented solutions (González 2011) and best management practices (Mora et al. 2011; Torres et al. 2011). Despite this recent progress, practical application of sustainable solutions is still limited in the context of Bogotá. Further efforts are needed to unify databases, link data analysis and modelling tools and simultaneously monitor the behaviour of the integrated system particularly during highly dynamic conditions (i.e. wet weather conditions) where underlying processes are more difficult to be appropriately captured and characterised.

CHAPTER 4:
CHARACTERISATION OF DRY WEATHER LOADS TO SEWERS

4.1 Introduction

Appropriate methods for quantifying temporal and spatial variations of inflows to sewer systems are a prerequisite to effective sewer system modelling. To contribute to this goal, an empirical generator of sub-catchment wastewater outflows, for use as inputs to dynamic simulations of the larger sewerage system, is developed and evaluated in this Chapter. The model is composed of four sub-models: (a) one that estimates mean wastewater flows and BOD, COD and TSS concentrations, (b) one that quantifies the mean transfer of flow and pollutant loads from the sanitary sewer system into the stormwater system via wrong connections, (c) one that represents mean infiltration flows into the sewer system, and (d) one that characterises flows and pollutant concentrations sub-daily variability around mean values.

4.2 Sub-model a: Mean flows and pollutant concentrations

For the Bogotá case a water user database provides estimates of water use from residential, industrial commercial, office-use and multi-user (i.e. apartment block) user types. The mean wastewater flows for the urban sub-catchments were estimated based on those aggregated consumption rates multiplied by a return factor of 0.84 (EAAB, personal communication), which is the reported value for Bogotá and the studied area. The water use data for 2006 were considered to be representative and used in the analysis. A stepwise regression analysis is used to establish a linear model that relates wastewater average flows and the number of users of each type in the sub-catchment. The stepwise analysis is basically a procedure for multilinear model selection in cases where there are a large number of potential explanatory variables and no clear underlying phenomena understanding on which to base the selection. It is a systematic method for adding and removing terms based on their statistical significance in a regression. The method begins with an initial model and then compares the explanatory power of incrementally more complex and simpler models. A Box-Cox transformation of both the flows and the user data is applied to make the data more normal distribution-like before applying the regression.

For concentrations, with only 29 out of 388 sewer sub-catchment outlets being gauged in Bogotá, there is a need for models that allow interpolation and extrapolation of the gauged concentrations to ungauged sites. This is achieved by identifying a relationship between sub-catchment properties (e.g. number and/or percentage of residential, industrial, commercial, office and multi users-related units, time lag, area, mean slope, hydraulic length and population) and the observed BOD, COD and TSS daily mean concentrations over the 29 gauged sub-catchment outflows. Two models were tested: (a) stepwise regression analysis and (b) a multi-objective search technique called “EPR MOGA”, which combines genetic algorithms and numerical nonlinear regression, and aims to simultaneously optimise model performance and parsimony (Giustolisi and Savic 2006; Giustolisi and Savic 2009). The selected models for flow and concentrations are then coupled with a stochastic term based on observed model errors.

4.3 Sub-model b: Average transfer of loads via wrong connections

Besides the already described data sets (see Chapter 3), different monitoring campaigns were carried out by EAAB (through local consultancy firms) in order to establish percentages of wrong connections from the sanitary into the stormwater system (i.e. number of properties that wrongly connected the wastewater outflow). For the Salitre catchment, a global value of 22% was established. (Mestra 2008) developed a GIS-based tool to identify the likelihood of cross-connections at the property level. Such a tool is used in this work in order to spatially distribute the global value of 22% among system sub-catchments. The main factors which have a relevant effect on the presence of cross-connections are urban densification processes, sewer system age, the physical gap between the storm water and wastewater systems, socioeconomic level, land use, pipe depth and distance between property and the wastewater and storm water systems, pipe material, road type and property type. Taking into account all these factors, Mestra’s tool labels the existence of wrong connections as being “low”, “medium”, and “high” likelihood. The global value of wrong connections is then linearly distributed based on the total number of properties with high likelihood labels. The tool was tested and the validated using a sub-catchment named Jaboque located in the Salitre sub-catchment in Bogotá (Mestra 2008).

4.4 Sub-model c: Mean infiltration flows into sewer systems

Regarding the infiltration flow, several studies have shown that the hydrology of urbanized areas can be very complex as the urban environment is highly heterogeneous in terms of land use, subsoil characteristics among other factors (Rodriguez et al. 2008), and hence infiltration flows can be difficult to estimate (Karpf and Krebs 2011). In this work a pragmatic approach to estimate those flows at each sewer sub-catchment in dry weather periods is used. Sewer infiltration, which is assumed uniformly distributed over all urban areas ($6x10^{-6}$ $m^3s^{-1}km\text{-}^2$), was calculated as the difference between (*a*) water supply distribution system leakages (which are about 2.1 m^3s^{-1}) and (*b*) the percolation flowing to underground layers (0.2 m^3s^{-1}) plus the annual evapotranspiration (0.03 m^3s^{-1} = 25% of the real evapotranspiration). According to these estimations, the average (in space) infiltration flow entering the sewer system corresponds to nearly 30% of the produced wastewater.

4.5 Sub-model d: Sub-daily variability

4.5.1 Deterministic component of sub-model

Modelling of DWF pollutant loads into sewer systems demands appropriate characterisation of the expected temporal and spatial variability of such system's inputs. According to the reviewed literature (see Chapter 2), inclusion of a stochastic component is needed if that variability is to be sufficiently characterised. Consequently, the model proposed in this Chapter consists of both a deterministic and a stochastic component. The two components are able to characterise DWF flow and quality determinants at gauged and ungauged sub-catchment outlets thus providing a mixed deterministic/stochastic representation of the time-space system variability. Following the recommendations of the Central European researchers simulation group (Langergraber et al. 2008) and common practice in integrated urban drainage modelling (ifak 2009), the diurnal patterns in input DWF data are modelled using Fourier series. Based on established procedure (e.g. as used in SWMM and InfoWorks), the DWF patterns are defined by a mean value and a set of hourly multiplier factors. Consequently, for each site and determinant, the first step is to normalise the observed data by dividing the mean. The normalised observed data are denoted by ${X_{i,j}}^{Obs}(t)$, where subscripts i and j refer to a particular site and a particular determinant respectively. The next step is to fit a Fourier series to ${X_{i,j}}^{Obs}(t)$ individually for each site and determinant. The trigonometric nth order Fourier series used has the form,

$$\mathbf{X_{i,j}(t) = a_{0_{i,j}} + \sum_{z=1}^{n} a_{z_{i,j}} cos(z \cdot t \cdot w_{i,j}) + b_{z_{i,j}} sin(z \cdot t \cdot w_{i,j})} \qquad \text{Equation 1}$$

where $a_{0_{i,j}}$ is a constant (with a value of 1 due to the applied normalisation), $a_{z_{i,j}}$ and $b_{z_{i,j}}$ are constants for each series order $(z = 1,2,\ldots,n)$, and $w_{i,j}$ is the fundamental frequency of the signal. By means of the *fit* function in Matlab, the optimisation of the a, b and w terms was performed adopting the coefficient of determination as the fit criteria. This procedure was applied using different values of n (i.e. 2, 3 and 4) aiming to achieve a good fit but to avoid over-fitting and thus reducing prediction uncertainty. For the Bogotá case, most of the wastewater quality measures were made using 3-hour composite samples (see Table 2). Where necessary, the Fourier series were fitted assuming that the 3-hour composite measurement applied to each of the three hours. The validity of using Fourier series in this way is assessed in the results sections.

$X_{i,j}(t)$ may be called a *global* time series as it aggregates the contributions from all five water use components (i.e. residential, industrial, commercial, office-use and multi-user). A method for disaggregating the global series is essential because the model must be used to estimate the determinants in sub-catchments which have water use combinations that are different from the 29 gauged sites. Hence the global time series are disaggregated into five time series: the component of $X_{i,j}(t)$ arising from each unit of residential, industrial, commercial, office-use and multi-user water use denoted as $R_j(t)$, $I_j(t)$, $C_j(t)$, $O_j(t)$ and $M_j(t)$ respectively. These five time-series are assumed to be common across all sites and therefore applicable also to ungauged sites,

$$\overline{X}_{i,j}(t) = re_i \cdot R_j(t) + in_i \cdot I_j(t) + co_i \cdot C_j(t) + of_i \cdot O_j(t) + mu_i \cdot M_j(t) \quad \text{Equation 2}$$

where for site i, re_i, in_i, co_i, of_i, mu_i are the known number of residential, industrial, commercial, office-use and multi-user units respectively and $\overline{X}_{i,j}(t)$ is the expected value of $X_{i,j}(t)$ for determinant j. Note that because the disaggregation is applied to normalised data and to concentrations rather than loads, this is an empirical analysis, rather than an analysis of contributions to the sub-catchment mass balance.

The estimation of the diurnal profiles $R_j(t)$, $I_j(t)$, $C_j(t)$, $O_j(t)$ and $M_j(t)$ can be approached in different ways. An elegant approach, in terms of minimising complexity and uncertainty, is to assume that each of the five profiles is a Fourier series such that the global series for any sub-catchment is the sum of five Fourier series (disaggregation A), In that case, 5 independent sets of a, b and w coefficients (totalling 40 coefficients per determinant if 3rd order Fourier series are used) have to be optimised. Alternatively, it is possible to make the less restrictive assumption that independent values of $R_j(t)$, $I_j(t)$, $C_j(t)$, $O_j(t)$ and $M_j(t)$ are estimated for each hour in the day hence not restricting the components to be Fourier series (disaggregation B, in which 120 coefficients have to be optimised per determinant). This model arguably lacks parsimony, however recall that (if all 29 gauged sites are used for model fitting) there are 696 observations per determinant. In either case, the coefficients are optimised by minimising the sum of the squared differences between the global values from Equation 1 and those from Equation 2,

$$f_j(t) = \sum_{t=1}^{24}\left[\sum_{i=1}^{I}\left(X_{i,j}(t) - \overline{X}_{i,j}(t)\right)^2\right] \quad \text{Equation 3}$$

where I is the number of gauged sites (29 in this case). For disaggregation A, the optimisation is done using the *mvregress* function in Matlab. For disaggregation B, the optimisation was implemented using the *lsqnonneg* function in Matlab, which constrains the solution to non-negative coefficient estimates. A comparison of disaggregation A and B performances can be found in the results section of this Chapter.

4.5.2 Stochastic component of sub-model

The stochastic error model aims to quantify the distribution of true values around the values estimated by Equation 2. A comprehensive error model could take into account all possible types of dependencies between errors including: (a) quantity and quality determinant autocorrelations, (b) inter-site cross-correlations, (c) inter-determinant cross-correlations and (d) inter-use (i.e. residential, industrial, commercial and official water uses) cross-correlations. In this work correlation types (b) and (d) are not looked at in detail: instead it is assumed that the errors between sub-catchments are independent and the water use related errors are perfectly correlated (i.e. each water use contributes proportionally to the total observed error). Including dependencies (b) and (d) more realistically is possible in principle, but would require more data than available in the case study and would lead to an extremely complex error model. Instead, the error model focuses on types (a) and (c), determinant autocorrelation and inter-determinant cross-correlations. The importance of including these correlations is evident considering that, for example, an underestimate of BOD from a sub-catchment is more likely than not to coincide with an underestimate of COD or TSS; and an underestimate of BOD at one time step is more likely than not to be preceded by another underestimate.

A model similar to the used by Matalas (1967) is used to generate synthetic sequences which adequately characterize and resemble historic error sequences. This model implies that the errors in normalised flows, BOD, COD and TSS are normally distributed and the serial correlation of each determinant may be described by a first-order linear autoregressive model. Multivariate synthetic sequences are defined as,

$$\begin{bmatrix} X'_{i,Q}(t) - \overline{X}_{i,Q}(t) \\ X'_{i,BOD}(t) - \overline{X}_{i,BOD}(t) \\ X'_{i,COD}(t) - \overline{X}_{i,COD}(t) \\ X'_{i,TSS}(t) - \overline{X}_{i,TSS}(t) \end{bmatrix} = A_i \begin{bmatrix} X'_{i,Q}(t-1) - \overline{X}_{i,Q}(t-1) \\ X'_{i,BOD}(t-1) - \overline{X}_{i,BOD}(t-1) \\ X'_{i,COD}(t-1) - \overline{X}_{i,COD}(t-1) \\ X'_{i,TSS}(t-1) - \overline{X}_{i,TSS}(t-1) \end{bmatrix} + B_i \begin{bmatrix} \varepsilon_Q \\ \varepsilon_{BOD} \\ \varepsilon_{COD} \\ \varepsilon_{TSS} \end{bmatrix}$$

Equation 4

where $X'_{i,j}(t)$ signifies a random realisation of determinant j at site i at time t, $\overline{X}_{i,j}(t)$ is the expected value of this determinant defined by Equation 2 as described above, and ε_j is a random normal variate with mean zero and standard deviation one. A_i and B_i are matrices (of size n-by-n, where n equals to the number determinants which in the case of this application is 4) whose elements are defined in such a way that the multivariate synthetic sequences generated will resemble the multivariate historic sequences in terms of error mean values, standard deviations, lag-one serial correlation for each determinant, and lag-zero cross-correlations between determinants (Matalas 1967). A_i and B_i are assumed to be independent of time, but they are dependent on the site, as described below. The elements of the matrixes A_i and B_i are given by,

$$A_i = M_{1_i} {M_{0_i}}^{-1} \qquad \text{Equation 5}$$

$$B_i {B_i}^T = M_{0_i} - M_{1_i} {M_{0_i}}^{-1} {M_{1_i}}^T \qquad \text{Equation 6}$$

where M_{0_i} is a matrix whose diagonal elements are the estimated error variances and whose off-diagonal elements are the estimated lag-zero error covariances. M_{1_i} is also a matrix whose elements are the estimated lag-one error covariances. The differences between the output of Equation 2 and the observed normalised data $\left(\overline{X}_{i,j}(t) - {X_{i,j}}^{Obs}(t)\right)$ are used as the observed sample of errors.

The need to generalise to ungauged sites requires simplifying assumptions about how the error properties vary between sites. First, for each determinant, it is assumed that each water use component contributes proportionally to the total observed error series. Using the residential component as an example, the standard deviation of the error in $R_j(t)$ is assumed equal to the standard deviation of the global error for determinant j at site i multiplied by $re_i \cdot R_j(t)/\overline{X}_{i,j}(t)$. Second, it is assumed that the error variance associated with a unit of each water component is uniform over all sites ($Variance_{R_j}$ for the residential component example), which can be estimated using a multivariate linear regression similar to Equation 2,

$$\begin{aligned} \mathrm{Variance}\left({X_{i,j}}^{Obs}(t) - \overline{X}_{i,j}(t)\right) \\ &= re_i \cdot \mathrm{Variance}_{R_j} + in_i \cdot \mathrm{Variance}_{I_j} + co_i \\ &\cdot \mathrm{Variance}_{C_j} + of_i \cdot \mathrm{Variance}_{O_j} + mu_i \cdot \mathrm{Variance}_{M_j} \end{aligned} \qquad \text{Equation 7}$$

Using the optimised coefficients of Equation 7, the variance of the error in $R_j(t)$ can then be calculated for any site, thus the global error for determinant j at any site is estimated from the observed errors across all gauged sites. And the third assumption required so that the covariance matrices can be generalised to ungauged sites, is that the cross-correlation and autocorrelation terms are uniform over all sites.

Applying Equation 4 for a particular ungauged sub-catchment produces a time series of errors for each of the four determinants which are then superimposed on the corresponding deterministic model output from Equation 2. If multiple realisations are simulated (i.e. multiple samples of $\varepsilon_{Q}, \varepsilon_{BOD}, \varepsilon_{COD}, \varepsilon_{TSS}$ in Equation 4, they represent the distribution of possible time series according to the model. If the model is adequate, in both its deterministic and stochastic components, the observed data will appear to be one realisation from the simulated distribution. Ideally the model would be assessed on each of the disaggregated components; however it is only possible to obtain measurements of the aggregated global response.

Although similar approaches to modelling error terms have been applied in other contexts, for example by Richardson (1981) to generate multi-variate time series of climate variables in gauged sites, this is its first application to stochastic generation of wastewater system model inputs for ungauged sewer areas as far as the author is aware.

4.6 Model assessment

The proposed modelling methods will result in probabilistic flow, concentration and loading predictions. The main assessment method is looking at the time series results to judge whether the observed values are samples from the modelled distributions. To help summarise results, the Nash-Sutcliffe efficiency adjusted for probabilistic predictions (Bulygina et al. 2009) is also used:

$$\mathrm{NSE}^{\mathrm{prob}} = \left\{1 - \frac{\sum_{t=1}^{T}(\mathrm{E}[\varepsilon_t] - x_t^0)^2}{\sum_{t=1}^{T}(x_t^0 - \mathrm{E}[x^0])^2}\right\} - \frac{\sum_{t=1}^{T}\mathrm{Var}[\varepsilon_t]}{\sum_{t=1}^{T}(x_t^0 - \mathrm{E}[x^0])^2} \qquad \text{Equation 8}$$

where ε_t is the sample of modelled values at time t, x_t^0 is the observed value at the same time, $\mathrm{Var}[\cdot]$ denotes variance, $\mathrm{E}[\cdot]$ denotes mathematical expectation (i.e. the mean value of the realisations), and T is the total number of time-steps being considered. The first part corresponds to the traditional NSE coefficient in which expected values are considered as predictors; and latter represents the variance whereby the higher predictor variance around the mean is, the less “effective” the prediction (Bulygina et al. 2009). The probabilistic NSE coefficient is used to help quantify model performance.

4.7 Results

This section firstly independently studies sub-models a and d because they contain the substantial scientific contribution and technical novelty of this Chapter. Secondly, this section also assesses to what extent coupled sub-models a to d allow characterisation of sub-catchment pollutant outflow diurnal dynamics. To do so, 5 gauged sites (out of the 13 sites described in Table 3) were selected: (i) site A corresponds to a combined sub-catchment outlet, (ii) sites C, F and G represent sanitary sub-catchment outlets and (iii) site D is a stormwater sub-catchment outlet.

4.7.1 Sub-model a: Identification of mean value models

Applying a stepwise regression analysis, as proposed in methods section, it was found that (at the 95% significance level) the sub-catchment outlet daily mean flow rate (in m^3s^{-1}) is linearly related to sub-catchment population, the number of residential, industrial, commercial, office-use and multi-use users with R^2 value of 0.89 (Figure 12a and Equation 9). However, linear models were incapable of usefully characterising mean concentrations, with optimised R^2 values of 0.14, 0.11 and 0.35 for BOD, COD and TSS respectively. Therefore the EPR MOGA-XL software (Giustolisi and Savic 2006; Giustolisi and Savic 2009) was used to explore if nonlinear models perform better. The expressions in Equation 10 and Equation 11 were identified for sub-catchment outflow BOD and TSS daily mean concentrations (in mg l-1). Equation 10 characterises observed BOD daily mean concentrations with an R^2 of 0.58 while Equation 11 characterises observed TSS with an R^2 of 0.61 (see Figure 12b and Figure 12c). COD is modelled using the linear relationship between BOD and COD as presented in Equation 12 and Figure 12c (no strong linear relationship exists for the couples BOD-TSS or COD-TSS). These models could be improved if additional data become available. However, they are considered the best models that can be identified from the currently available datasets.

$$Q = 4.38x10^{-6} \cdot re_i + 5.71x10^{-5} \cdot in + 8.60x10^{-6} \cdot co_i + 0.00027 \cdot of_i + 6.95x10^{-6} \cdot mu_i M$$ Equation 9

$$BOD = 0.0048 \cdot r^2 \cdot i^{0.5} \cdot m^{0.5} + 23992 \frac{{re_i}^{0.5} \cdot m^{0.5} \cdot L^{0.5}}{H^{1.5}} + 1.51 \frac{{re_i}^{0.5} \cdot r}{K^{0.5}}$$ Equation 10

$$TSS = 6.33 \cdot \frac{i'^{0.5} \cdot (K)^{1.5} \cdot L}{{co_i}^2} + 0.00074 \cdot in_i \cdot co_i + 2.33 \cdot {re_i}^{0.5}$$ Equation 11

$$CDO = 2.05 \cdot BOD$$ Equation 12

where r, i and m are the percentage of residential, industrial and multi-users units; L is the sub-catchment hydraulic length (in m); i' is the percentage of industrial water use consumption; K is the sub-catchment travel time (in seconds) and H is the number of inhabitants. The performances of Equation 9 to Equation 12 for other case studies are unknown and further research is needed to identify if the significant explanatory variables vary from one sewer system to the next. For this, averaging EPR-based models over multiple sewer systems has proved to be useful (Savic et al. 2009).

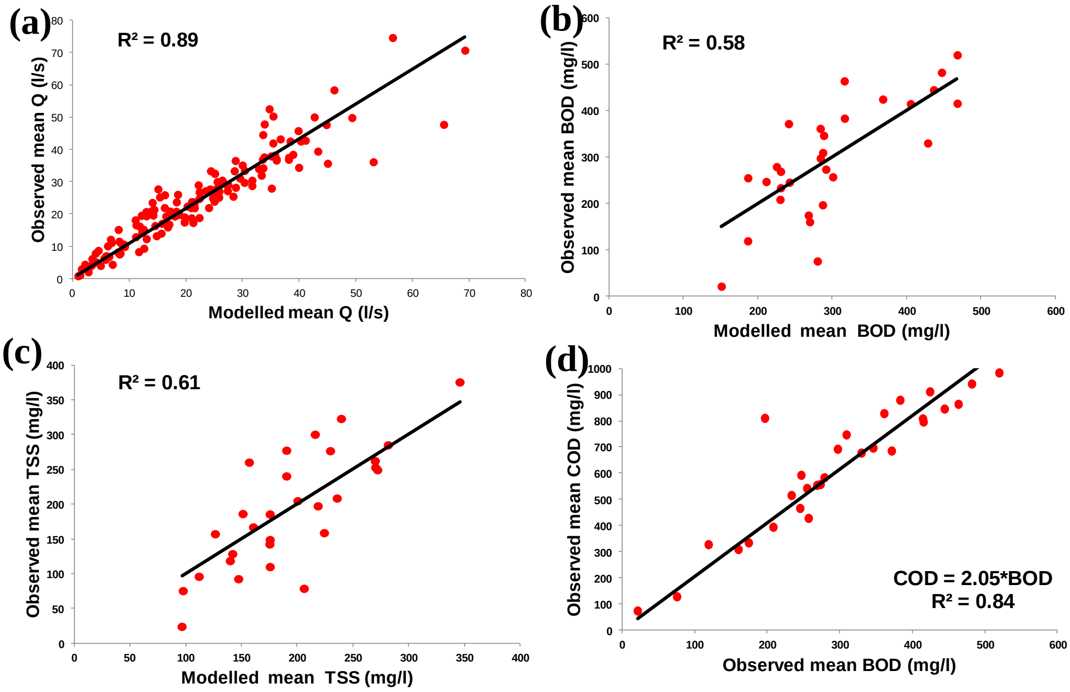


Figure 12. Comparison between observed and modelled data: (a) flow; (b) BOD; (c) TSS daily mean concentrations; (d) linear relationship between observed BOD and COD daily mean concentrations

4.7.2 Sub-model d: Sub-daily variability

Deterministic component assessment

As mentioned in the preceding section of this Chapter, the first step in the proposed time series generation scheme is to fit an nth order Fourier series to the measured data. With guidance from previous studies (Carstensen et al. 1998; Bechmann et al. 1999; Langergraber et al. 2008; Mannina and Viviani 2009; Alex et al. 2011), the performance of 2nd, 3rd, and 4th order Fourier series are compared. In general, it was found that 3rd order Fourier series gave better fits to the measurements than 2nd order series, and no further improvement is achieved if 4th order series are used (Table 4). Consequently, 3rd order series were selected to represent the deterministic component when modelling DWF patterns.

Table 4. Goodness of fit to measured data when using 2nd, 3rd and 4th order Fourier series[1]

	2nd order Fourier Series		3rd order Fourier Series		4th order Fourier Series	
	Mean R^2	St.dev. R^2	Mean R^2	St.dev. R^2	Mean R^2	St.dev. R^2
Flow	0.73	0.18	0.80	0.16	0.78	0.18
BOD	0.79	0.08	0.85	0.06	0.86	0.06
COD	0.79	0.12	0.86	0.10	0.84	0.11
TSS	0.74	0.12	0.81	0.08	0.80	0.11

[1] Mean and standard deviation are calculated over the 29 sites

Sub-catchment ID 13 (150 ha with about 88%, 6% and 6% of residential, industrial and commercial water uses respectively, see Figure 4 and Figure 5 for more details) is used to exemplify the 3rd order Fourier series results, illustrating that the fitted Fourier series usefully characterise the observed 3-hour average data. Table 5 summarises the optimised values for each of the coefficients of Equation 1 including the corresponding 95% confidence interval when fitting flow, BOD, COD and TSS respectively. It can be concluded from Table 5 that not only the coefficients but the fundamental frequency is different for each determinant thus reinforcing the need to independently model each.

Using sub-catchment ID 10 (96% residential water consumption, representative of an average sub-catchment in Bogotá, see Figure 4 and Figure 5 for more details) as an example, Figure 13 also illustrates the general ability of the 3rd order Fourier series to represent the observed wastewater quality dynamics, and also shows the sensitivity results of using 3-hour composite sample measurements rather than hourly measurements, as was necessary in most sub-catchments. Figure 13 supports the previous conclusion that Fourier series are able to characterise the observed hourly data, and the results have limited sensitivity to the use of composite samples. Industrial areas, which have a more dynamic water use profile, are likely to exhibit greater sensitivity to the use of composite samples, however this could not be tested because of the lack of non-composited hourly data.

A possible way to improve the parsimony of the model would be to assume that TSS/BOD and TSS/COD are constant values (Mannina and Viviani 2009) and applying Equation 2 and Equation 4 only to TSS. However, it was found that there is large intra-site hourly variability of the ratios TSS/BOD and TSS/COD, and also considerable inter-site variability within the Bogotá sewer sub-catchments. For example, the inter-site mean and standard deviation of the TSS/BOD intra-site mean values are 0.9 and 0.9 respectively, and the corresponding values of TSS/COD are 0.4 and 0.4. The results show the limitations of assuming fixed linear relationships between water quality determinants, and justify the decision to use separate BOD, COD and TSS time series models in this case.

Once 3rd order Fourier series have been fitted to available data from gauged sub-catchments, those global time series are disaggregated into five time series ($R_j(t)$, $I_j(t)$, $C_j(t)$, $O_j(t)$ and $M_j(t)$ respectively) as described above. Figure 14 presents results of the disaggregated time series for TSS as an example, using both disaggregation methods A and B. It presents the optimised values when including data from all 29 gauged sites in the fitting and the estimated 95% confidence bounds which were obtained applying a bootstrap analysis which consisted of optimising over 1,000 random re-samples of the 29 sites (Efron 1979). Although each of the random re-samples comprised data of 29 sites, a particular gauged site can be considered more than once within a particular sample. Figure 14 shows that there is large uncertainty in the coefficients using both disaggregation methods, caused by the limitations of the model in fitting the data and also the relatively small number of data points per model. However, disaggregation A led to reduced uncertainty particularly for residential and industrial contributions (see Figure 14). Consequently, disaggregation A is preferred over B and used for the rest of this Chapter. Figure 15 presents R^2 values when comparing $X_{i,j}(t)$ and $\overline{X}_{i,j}(t)$ for each site and determinant to help quantify the overall efficiency of the disaggregation procedure A. Due to the large uncertainty in the coefficient estimates, R^2 values can be poor for some sites and determinants. However there is not any particular site or determinant that performs consistently badly (see Figure 15). The implications of the uncertainty for usefulness of model outputs will be shown later in the validation tests.

Table 5. Comparison of 3rd order Fourier series parameters in a particular urban sub-catchment (i.e. site ID 13: 149 ha, nearly 88% of residential water use, 6% of industrial use and 6% of commercial use) for flow, BOD, COD and TSS

Coefficient	Flow*	BOD*	COD*	TSS*
$a_{0_{13,j}}$	1.0 (0.92, 1.09)	1.04 (0.93, 1.153)	1.05 (0.99, 1.117)	1.10 (0.96, 1.23)
$a_{1_{13,j}}$	-0.30 (-0.53, -0.08)	-0.31 (-0.63, 0.002668)	-0.22 (-0.37, -0.0805)	-0.37 (-0.56, -0.17)
$b_{1_{13,j}}$	-0.28 (-0.49, -0.073)	0.41 (0.18, 0.6305)	0.30 (-0.40, -0.1787)	-0.13 (-0.34, 0.08)
$a_{2_{13,j}}$	-0.001 (-0.12, 0.10)	0.001 (-0.20, 0.21)	0.10 (0.01, 0.19)	0.52 (0.30, 0.74)
$b_{2_{13,j}}$	-0.07 (-0.21, 0.07)	0.15 (-0.05, 0.35)	-0.05 (-0.13, 0.04)	-0.11 (-0.54, 0.32)
$a_{3_{13,j}}$	-0.05 (-0.17, 0.08)	-0.04 (-0.35, 0.26)	-0.12 (-0.20, -0.03)	-0.28 (-0.63, 0.07)
$b_{3_{13,j}}$	0.04 (-0.15, 0.21)	-0.16 (-0.35, 0.04)	0.041 (-0.14, 0.22)	-0.24 (-0.60, 0.12)
$w_{13,j}$	0.26 (0.21, 0.30)	-0.28 (-0.33, -0.23)	0.31 (0.27, 0.34)	0.34 (0.31, 0.38)

*with 95% confidence bounds

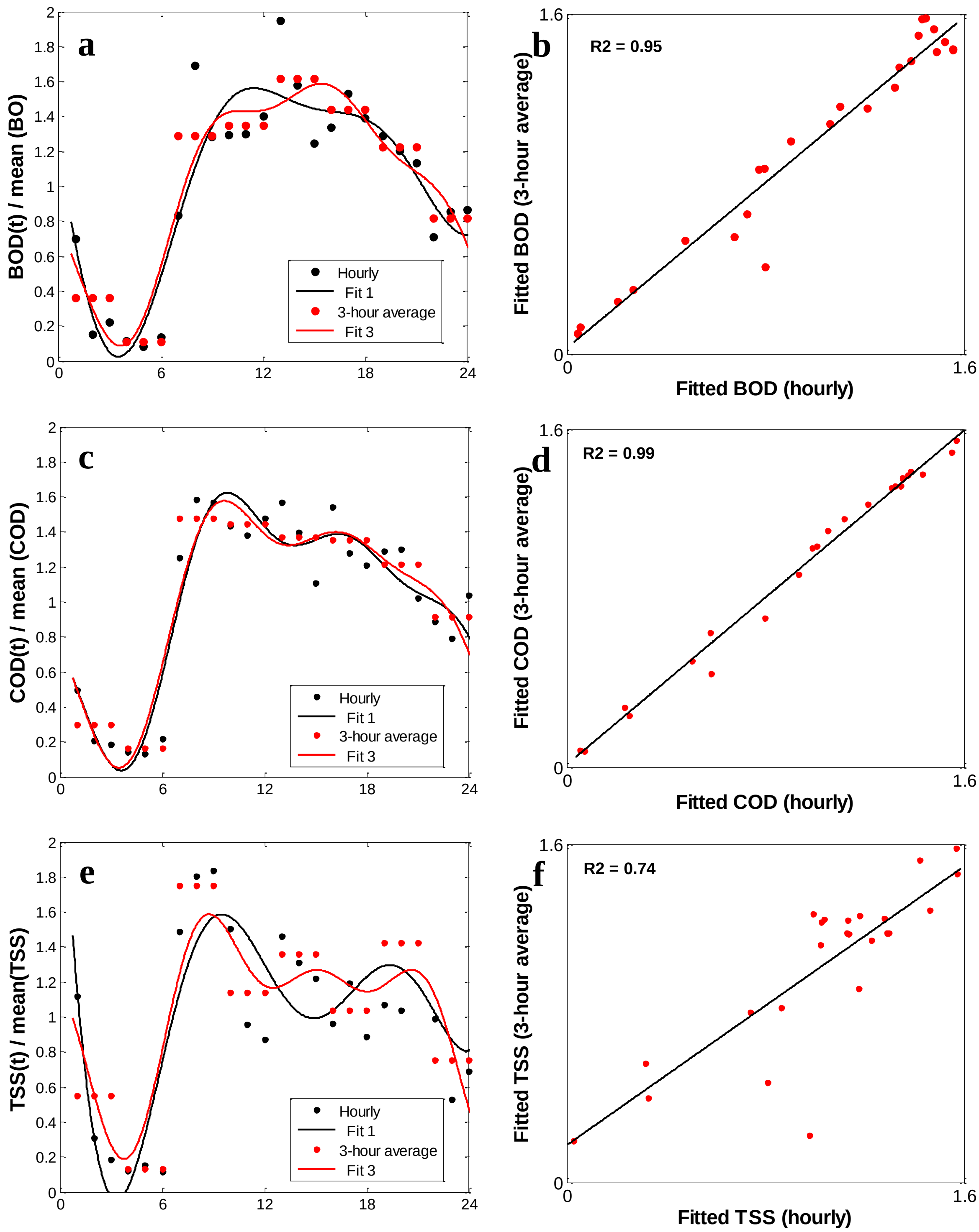


Figure 13. Comparison of 3rd order Fourier series fitted to observed hourly (see Fit 1) and 3-hour average (see Fit 3) for (a) BOD, (c) COD and (e) TSS values at sub-catchment ID 10 outlet. Comparison of hourly data estimated using Fit 1 and Fit 3 for (b) BOD, (d) COD and (f) TSS

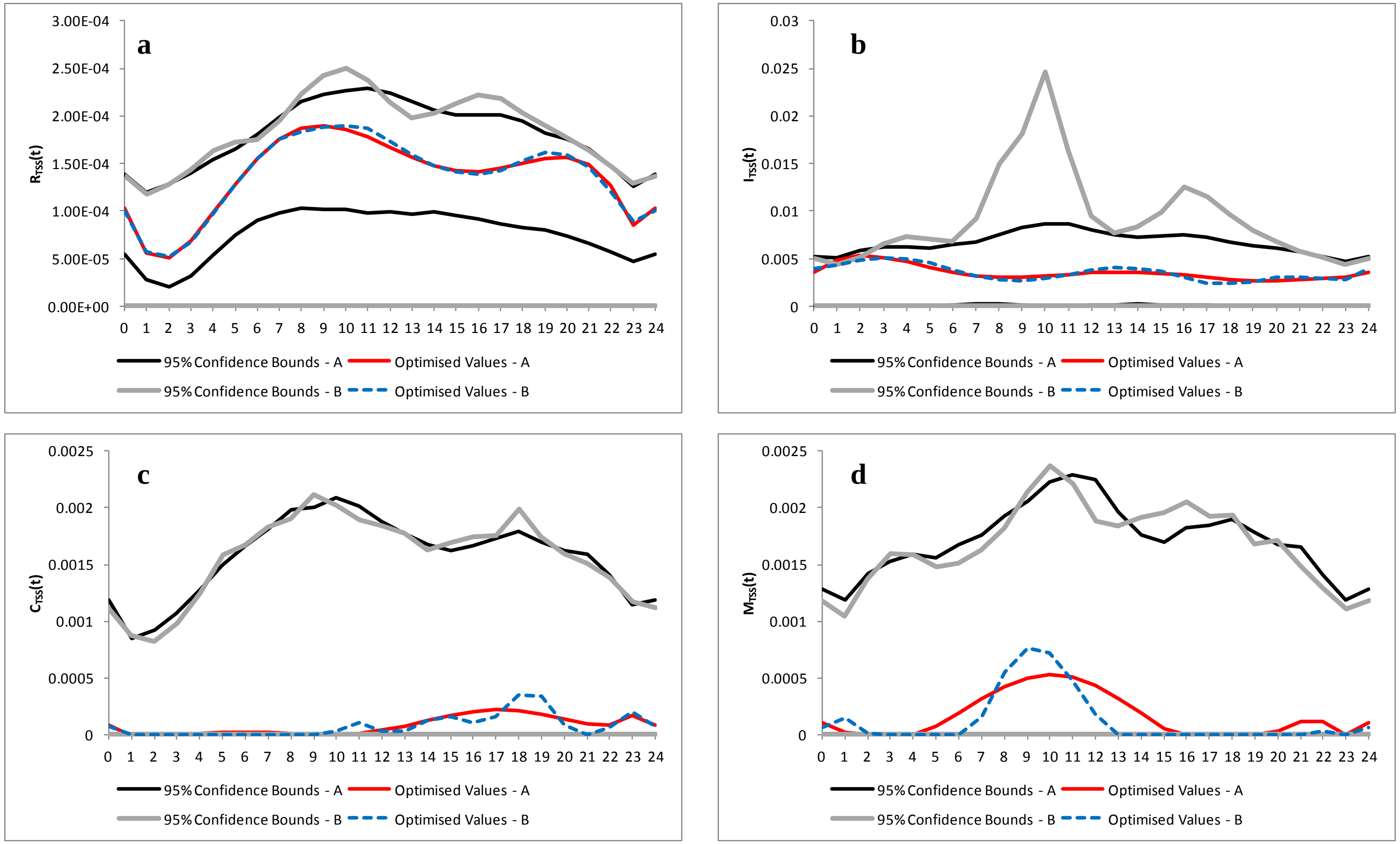


Figure 14. Comparison of estimates for $\mathbf{R_{TSS}(t)}$, $\mathbf{I_{TSS}(t)}$, $\mathbf{C_{TSS}(t)}$, and $\mathbf{M_{TSS}(t)}$ and their related uncertainty when fitted to the 29 gauged sub-catchments simultaneously. Results shown for disaggregation methods A and B.

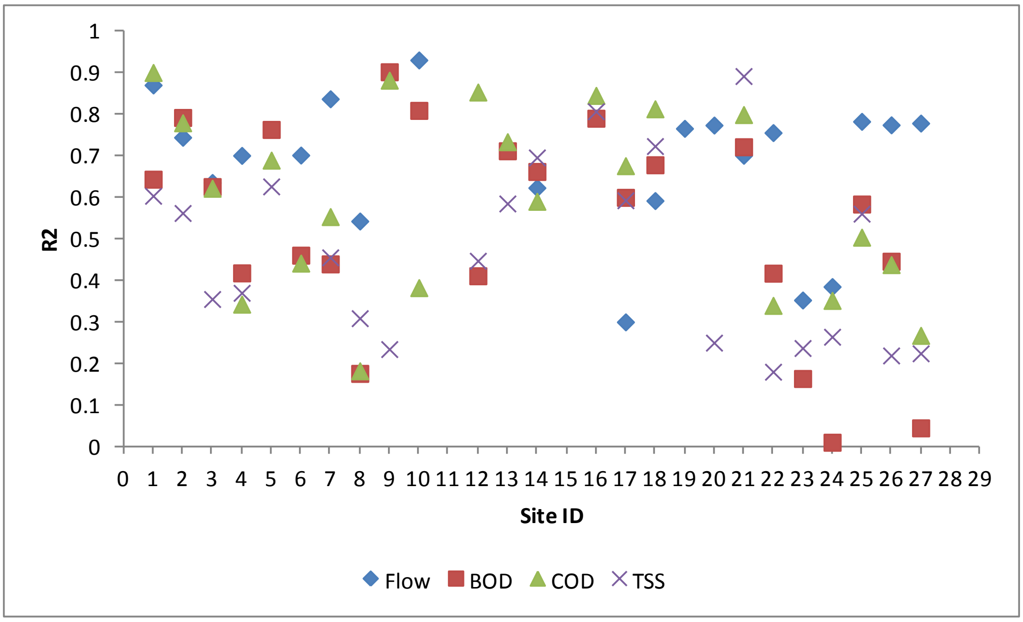


Figure 15. R^2 values when comparing $\mathbf{X_{i,j}(t)}$ and $\mathbf{\bar{X}_{i,j}(t)}$ for each site and determinant to help quantify the overall efficiency of the disaggregation procedure A

Error analysis

The observed errors should ideally be consistent with the main assumptions employed in the model: the errors are described by a zero-mean multi-variate normal distribution and a 1st order linear autoregressive model, and the error properties are stationary over time and over sites. For this error analysis, all 29 sites were used in the model fitting.

The normal distribution assumption

To graphically illustrate the validity of assuming a normal distribution, the mean and skewness coefficient of the observed errors were calculated for each site and determinant (see Figure 16). Besides this, the Lilliefors test (Lilliefors 1967) of the default null hypothesis that there are not significant differences between the statistics of the observed errors and those of a normal distribution at the 5% significance level and 95% confidence level was performed. For flow errors, in 7 out of the 29 sub-catchments, there is enough evidence to reject the null hypothesis. The same was observed in 6 and 2 sites for COD and TSS errors respectively. This implies that there is little information in the data which could be used to improve the deterministic or stochastic models. However, the validation tests presented later in this Chapter show that the general variability of the observations is captured by the model and so it was preferred not to pursue a more complicated model. To further support the normal distribution assumption, Figure 24 and Figure 26 show examples of near-normally distributed errors for complementary case studies.

1^{st} order linear autoregression assumption

The correlation coefficients were calculated for each observed error series (all determinants at all measured sites) for lags up to 5 hours. As suggested by Richardson (1981), these serial correlation coefficients were compared to a first-order autoregressive model given by,

$$\rho_K = \rho_1{}^K \quad \text{Equation 13}$$

Where ρ_K is the serial correlation for lag K (in hours) and ρ_1 is the lag 1 serial correlation coefficient from the error series. Figure 17 shows a representative comparison for one of the gauged sites, indicating that there is consistent lag-2 autocorrelation, not captured by the autoregressive model. Similar results were obtained for the rest of the validation sites. A particular problem may exist for BOD and COD: the observed lag-2 autocorrelations were significantly different from the modelled values at 15 and 14 sites for BOD and COD respectively. The reason of this lag-2 autocorrelation is presumed to be the use of 3-hour average values for the hourly estimation as illustrated in Figure 21 and Figure 22. The practical significance of this problem will be reviewed in the performance assessment.

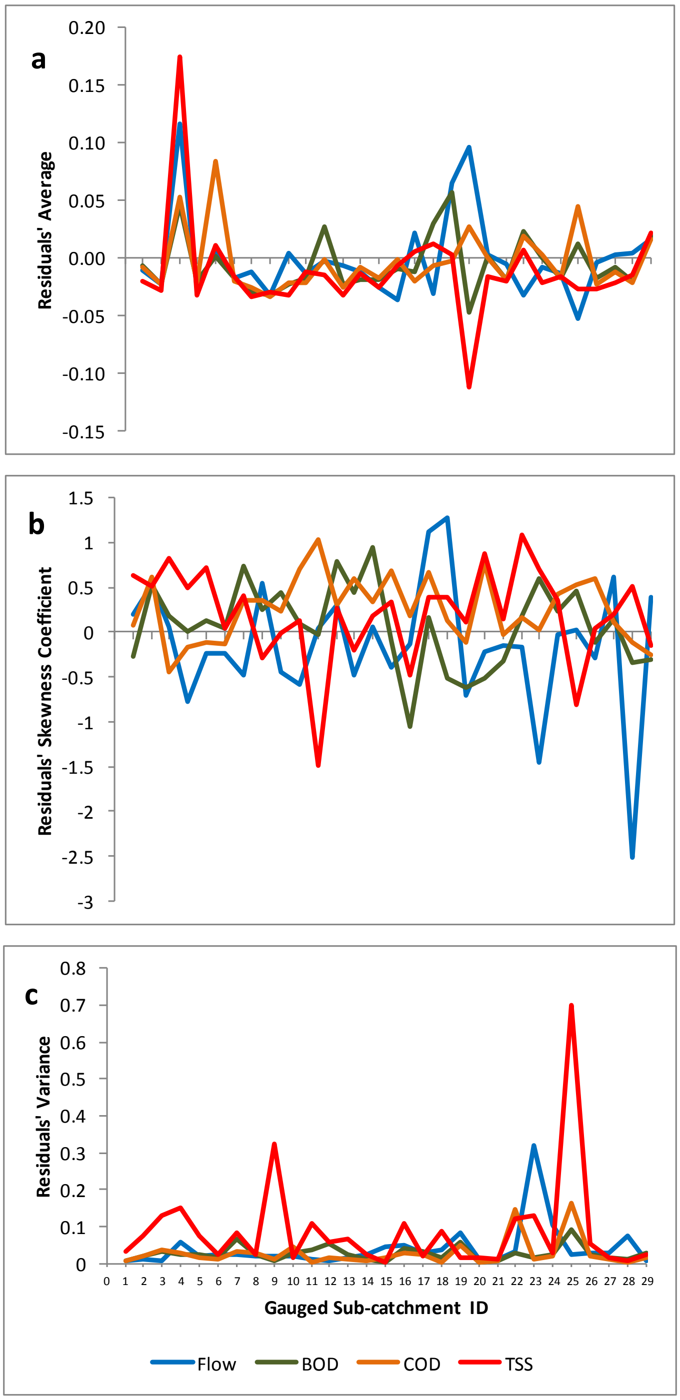


Figure 16. (a) Mean, (b) Skewness coefficient and (c) Variance of Flow, BOD, COD and TSS errors. Errors were calculated for the normalised time series from the corresponding 3rd order Fourier fitted series

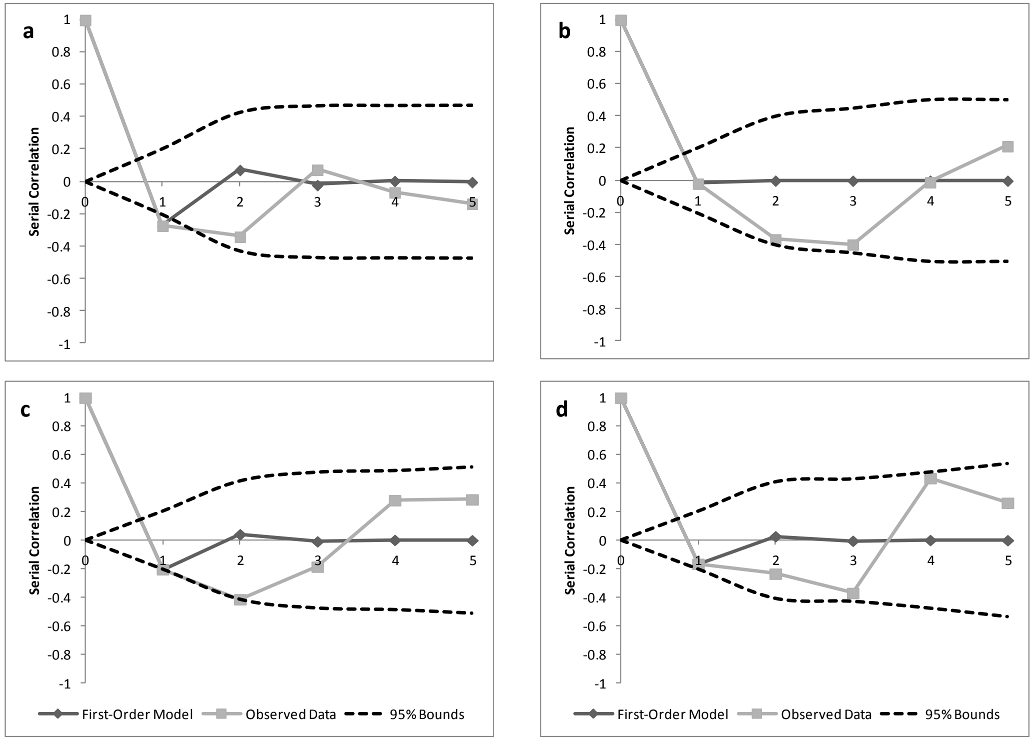


Figure 17. Serial correlations coefficients (for one of the validation sites -ID 13- where water consumption is about 88% residential) of (a) flow, (b) BOD, (c) COD and (d) TSS compared to a first-order linear model given by $\boldsymbol{\rho_K = \rho_1{}^K}$. Additionally, upper and lower bounds for autocorrelation with significance level 5% were estimated using the Bartlett's formula

The inter-determinant dependence model

The cross-correlation analyses shown in Figure 18 illustrate the overall significance of the inter-determinant dependencies. If the inter-dependence model for flow, BOD, COD and TSS is valid, the observed distributions of errors will be close to the bivariate normal distributions assumed in Equation 4. Figure 19 presents plots of the observed errors for each possible pair of these determinants (i.e. flow-BOD, flow-COD, flow-TSS, BOD-COD, BOD-TSS and COD-TSS) including data from all 29 sites. Although only a visual analysis, the figure shows there is no clear evidence that the observed errors diverge from the assumed distribution. To confirm this, using the Henze-Zirkler test (Henze and Zirkler 1990) it was found that the errors at 21 of the 29 the gauged sub-catchments correspond to a bivariate normal distribution at a significance level of 95% for the flow-BOD pair. For the pairs flow-COD, flow-TSS, BOD-TSS and COD-TSS: 23, 20, 28, 28 and 24 sites appear to correspond to the bivariate normal assumption. Besides this, there is evidence that the correlation coefficient varies between sites (Figure 20), contrary to the uniformity assumption. However, it was not possible to identify a systematic variation between sites, so the assumption is considered pragmatic.

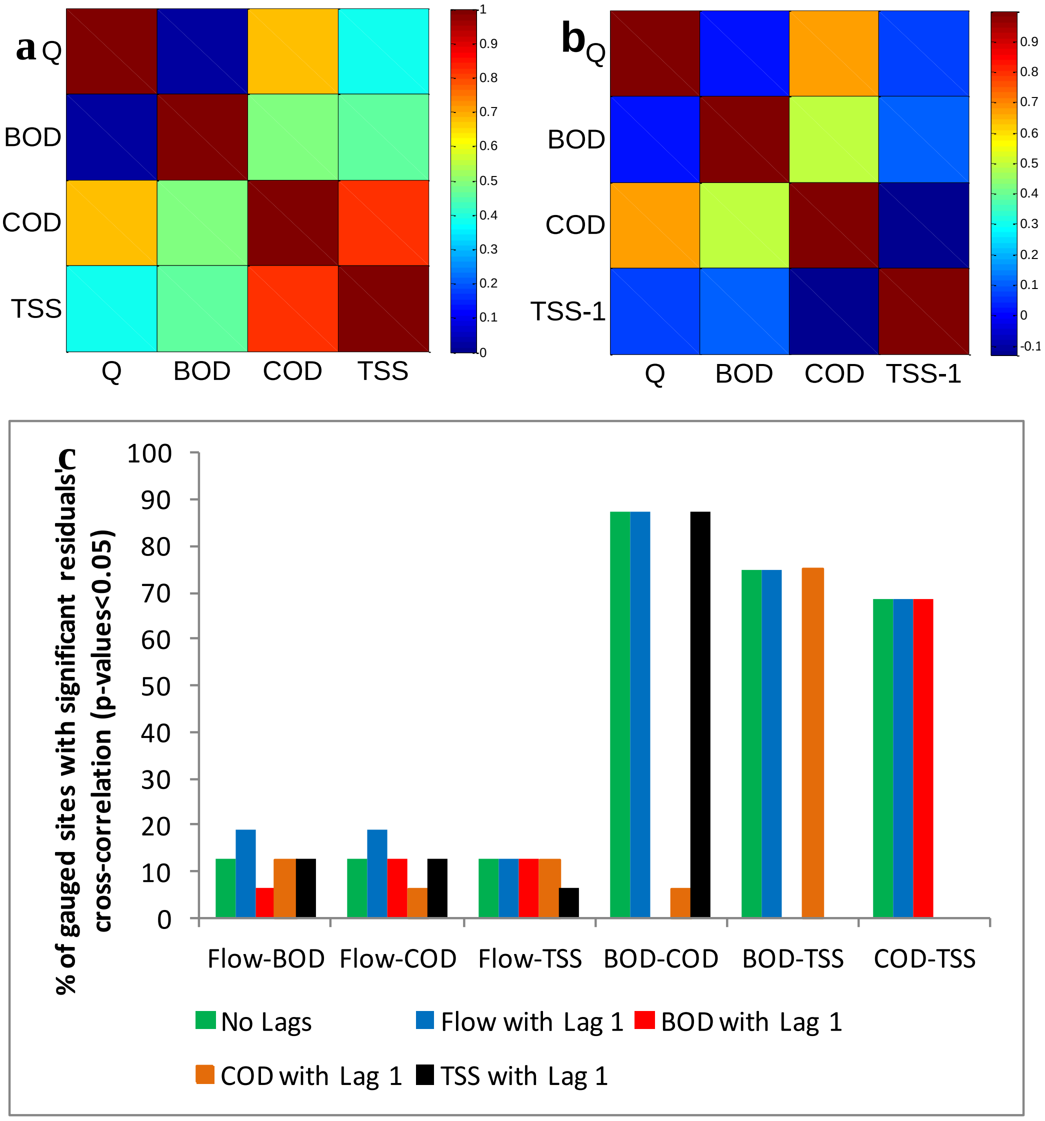


Figure 18. (a) Cross-correlation coefficients between the errors of flow, BOD, COD and TSS for a particular site (site ID 1); (b) Cross-correlation coefficients between the lag 1 errors of flow, BOD, COD and TSS for a particular site (site ID 1); (c) % of gauged sites with significant a cross-correlation between errors for lag 0 and lag 1. Figures (a) and (b) follow guideline 3 made by (Kelleher and Wagener 2011)

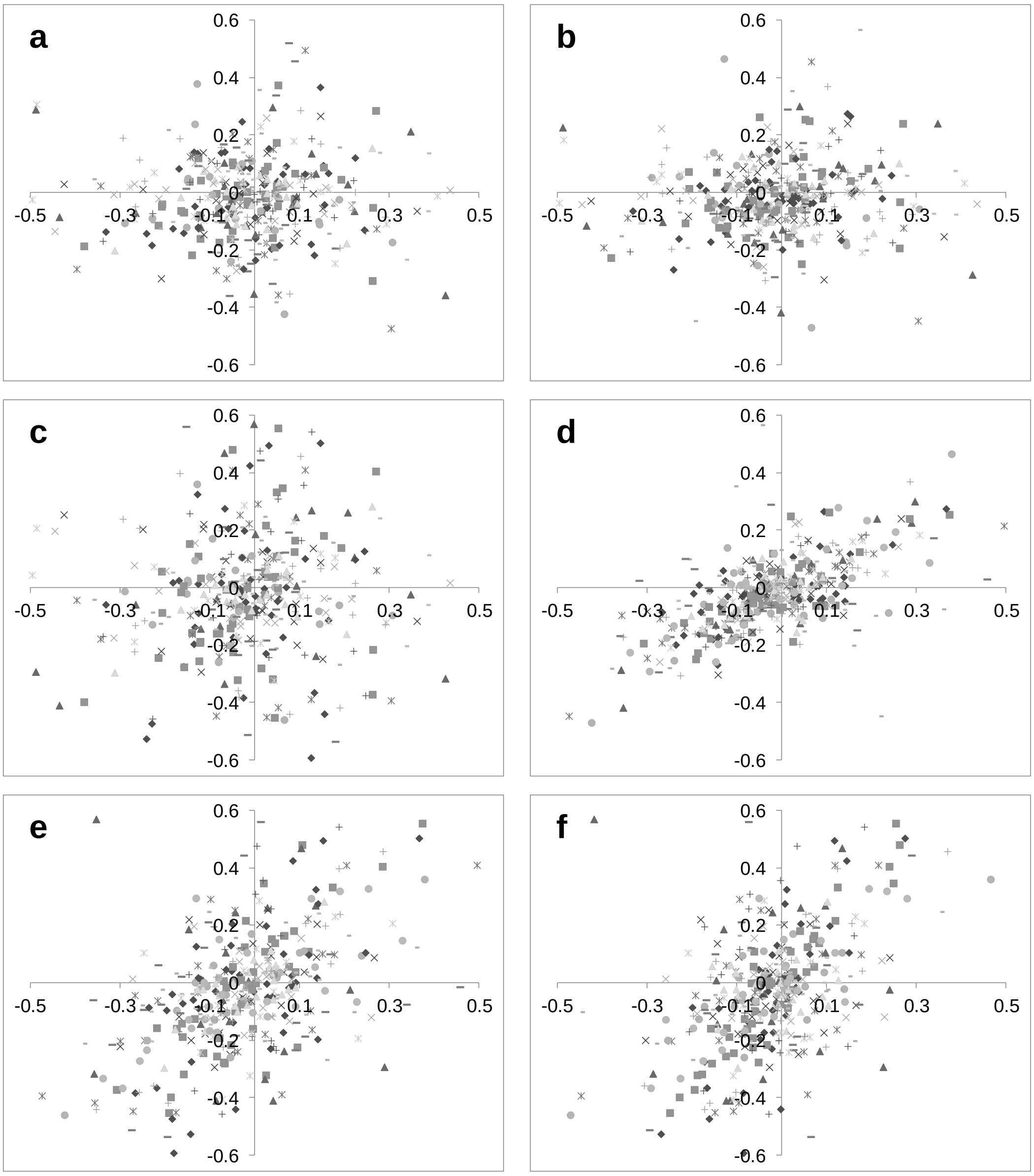


Figure 19. Observed errors for pairs: (a) flow-BOD, (b) flow-COD, (c) flow-TSS, (d) BOD-COD, (e) BOD-TSS and (f) COD-TSS. Different symbols correspond to the 29 different gauged sites

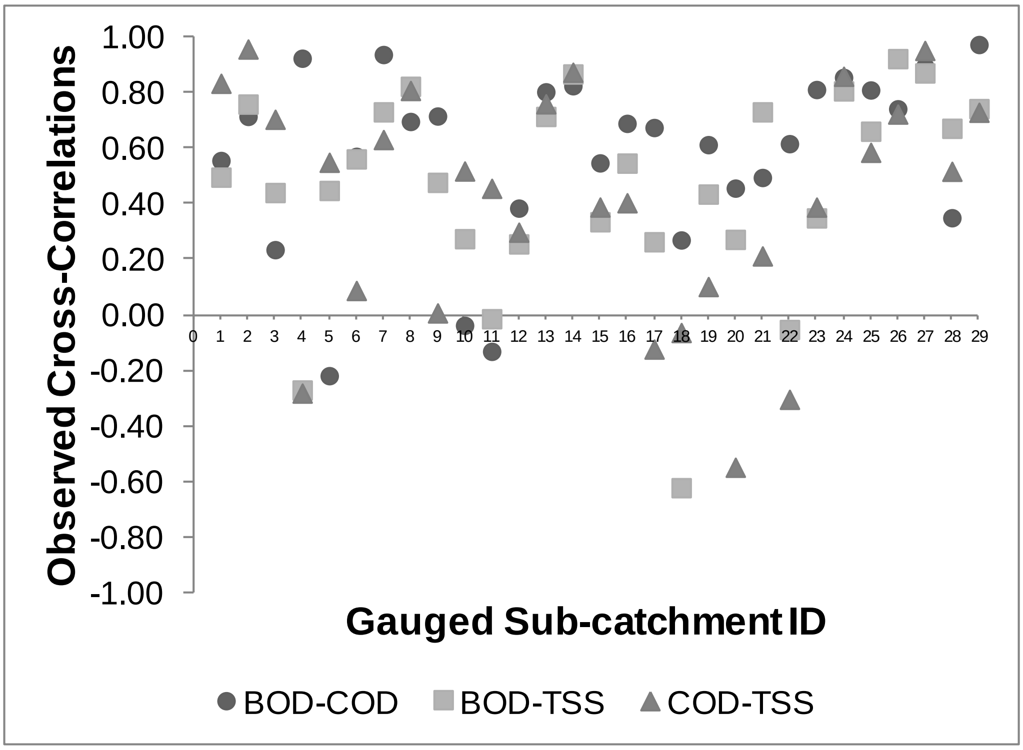


Figure 20. Observed error correlations for pairs BOD-COD, BOD-TSS and COD-TSS

Validation of performance on ungauged areas

In order to investigate the ability of the proposed stochastic generator to characterise ungauged areas, a jack-knife procedure (Khalil et al. 2011) was used. Within this procedure, each of the 29 measured sites was used as a test site and temporarily assumed to be ungauged. Then, the model was fitted using only data from the other 28 gauged sites and predictions made at the test site. The main objective of this analysis was to identify to what extent the modelled uncertainty bounds include the observed values for flows, concentrations and pollutant loads, and to judge how well individual realisations represent the observed variability.

For each of the sites a total of 1,000 error realisations were carried out (this number of realisations was sufficient to appropriately characterise error variability), allowing the estimation of 95% confidence bounds. The jack-knife validation results for two contrasting sub-catchments (site ID 6 which is a 58 ha area with about 70% and 30% of residential and commercial water use respectively; and site

ID 13) are in Figure 21 and Figure 22. These figures present the following validation results (i.e. observed data is not used to estimate neither the deterministic nor the stochastic components of model): (*a*) the observed values, (*b*) the modelled 95% confidence bounds, (*c*) the deterministic component of the model, and (*d*) three examples of time series (superimposing the stochastic component to the deterministic component of the model) generated by the model. A relatively high proportion of all observed values of flow, determinant concentrations and loads are included within the modelled bounds (see Table 6).

The previously noted evidence that lag-2 autocorrelations might affect the cumulative loads over, say, a 3-hour period, which might be relevant for WWTP operation. However, the 3-hour loads were slightly better represented by the models than the instantaneous loads (see Table 6). This supports the use of the use of the 1st order autoregression error model.

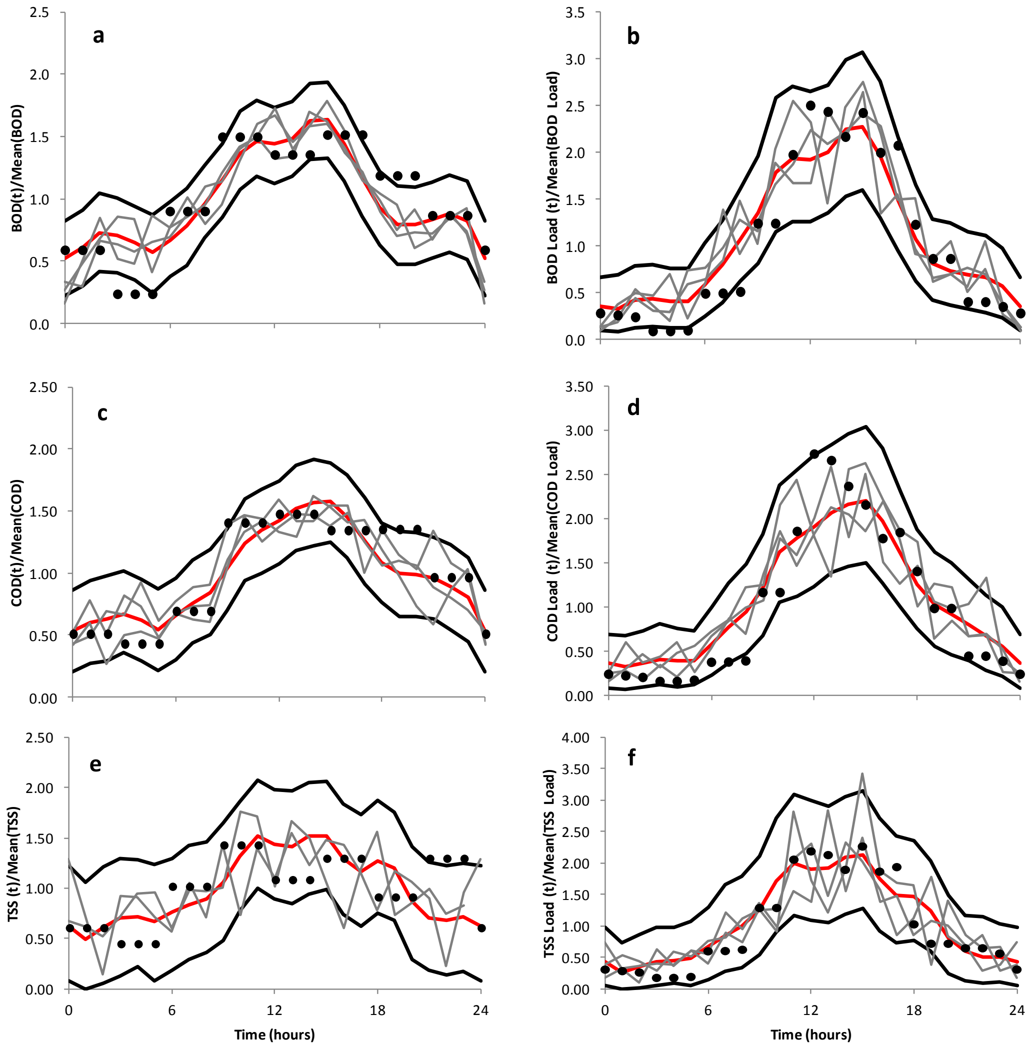


Figure 21. Observations and jack-knife validation results for site ID 6: (a) BOD concentration, (b) BOD load, (c) COD concentration, (d) COD load, (e) TSS concentration and (f) TSS load. Each sub-plot includes: Observed data (black dots), the deterministic model output (red line), 3 representative error realisations (grey lines), and 95% confidence bounds derived from 1,000 realisations

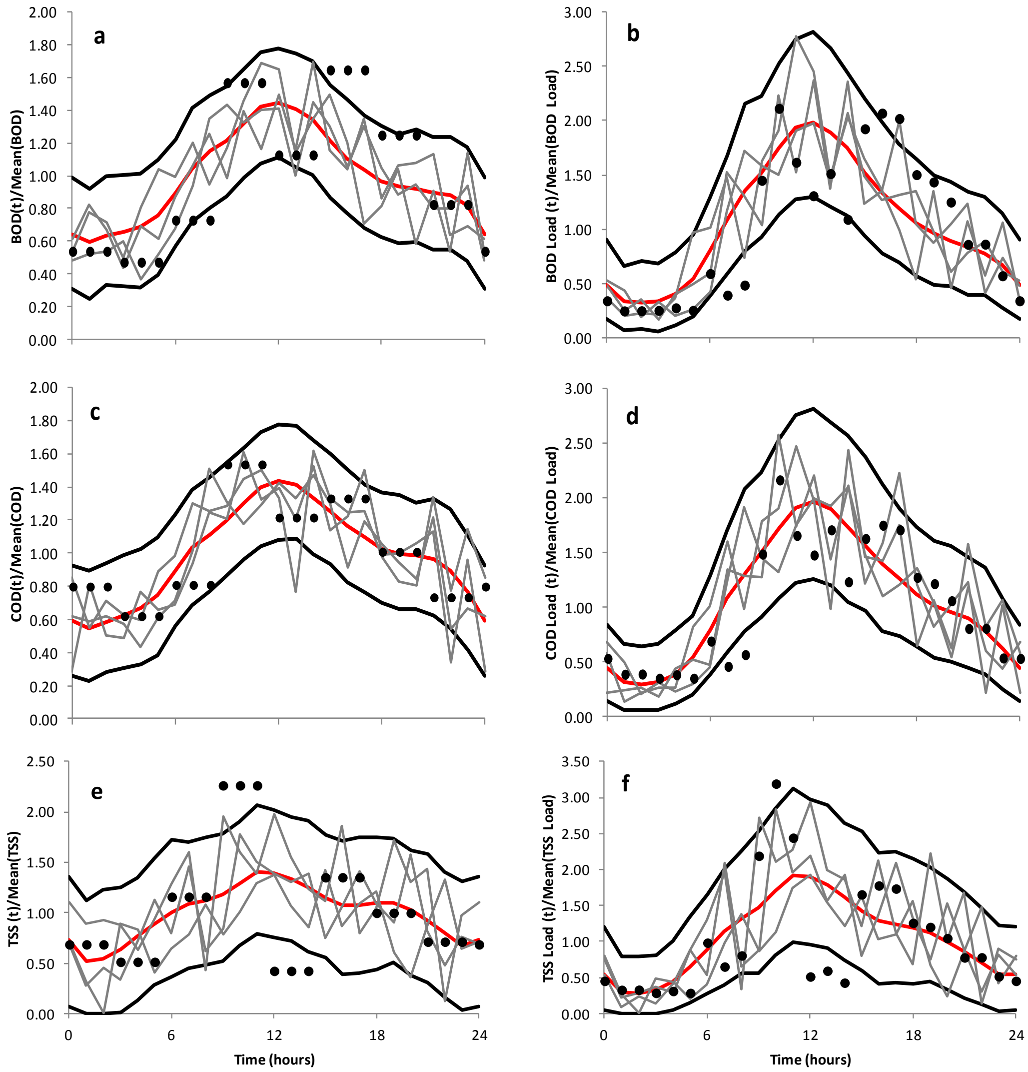


Figure 22. Observations and jack-knife validation results for site ID 13: (a) BOD concentration, (b) BOD load, (c) COD concentration, (d) COD load, (e) TSS concentration and (f) TSS load. Each sub-plot includes: Observed data (black dots), the deterministic model output (red line), 3 representative error realisations (grey lines), and 95% confidence bounds derived from 1,000 realisations

Table 6. % of observed values lying within the simulated 95% intervals at ungauged sites using jack-knife validation procedure[2] (*In percentage)

	Determinant		Load		Cumulative Loads (3 hours)	
	Average*	St.dev.*	Average*	St.dev.*	Average*	St.dev.*
Flow	79	20	N/A	N/A	N/A	N/A
BOD	73	19	77	19	83	19
COD	80	18	80	16	86	17
TSS	96	8	96	5	98	4

[2] Mean and standard deviation are calculated over the 29 sites

Model transferability to other case studies

Three key assumptions made in the model are that: the global signals of flow and quality can be represented by a Fourier series; the random variations around the signals are normally distributed; and the serial dependence of the errors is approximated by a 1st order linear autoregressive model. Results have suggested that these are reasonable assumptions in the Bogotá case. This section evaluates if these assumptions apply equally to different types of sub-catchments with higher resolution. Two case studies are used: the El Virrey sub-catchment in Bogotá (data from this sub-catchment were not used above); and a sub-catchment within the urban drainage system of Linz, Austria. This evaluation does not cover the transferability of the ungauged site model, for which additional extensive spatial data sets would be required.

In the El Virrey sub-catchment the water consumption is residential (about 40%) and commercial (about 60%). Flow and quality were monitored using 3 automatic samplers coupled with one ultrasonic level sensor and one multiparametric sonde between August 2000 and August 2001 at one sub-catchment outlet named Caracas 77, with a catchment area of 85 ha (Uniandes 2001) (Figure 23). Using the 1-minute flow data over 51 dry days, the errors from a 3rd order Fourier series model were near-normally distributed (Figure 24a) with mean and skewness coefficient of

0.0 and 0.22 respectively. The correlation coefficients were calculated for lags up to 5 hours, showing a close approximation to the 1st order autoregressive model in Equation 13 (Figure 24b).

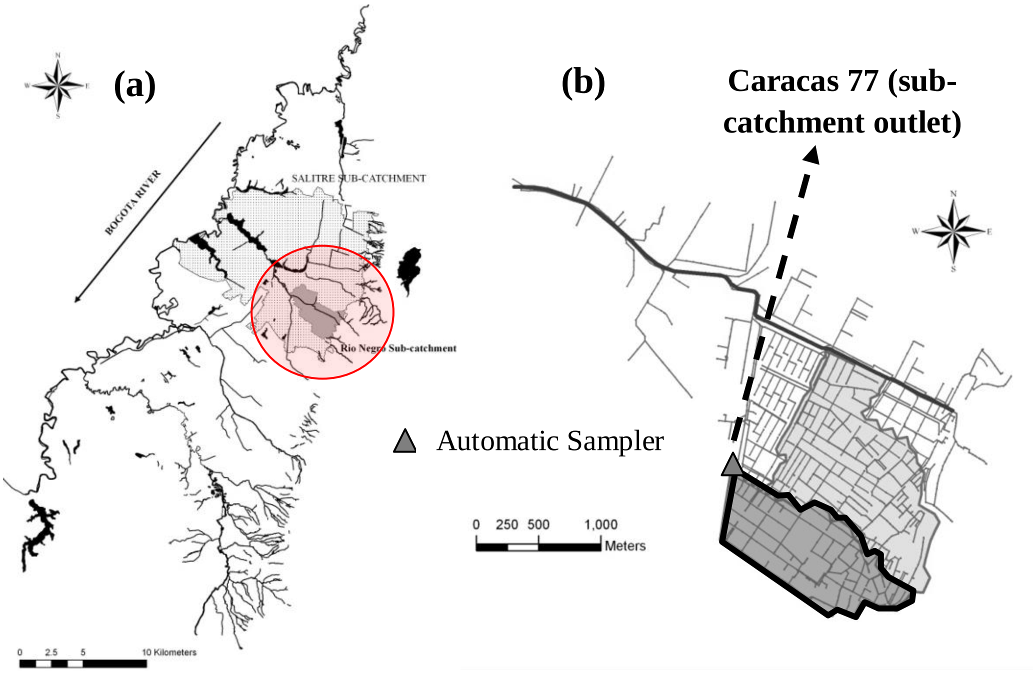


Figure 23. The el Virrey sub-catchment: (a) location within Bogotá's sewer system, (b) location of the monitoring station

Linz is the third largest city of Austria and capital of the state of Upper Austria. A combined sewer system drains the city of Linz and 39 surrounding communities to the WWTP Linz-Asten, with 300,000 population equivalent connected (Achleitner et al. 2009). These communities have an important influence especially in sub-catchments 4 and 5 (Figure 25). The system consists of six main sub-catchments (as presented in Figure 25 and Table 7). The sub-catchment 1, which is a mixture of rural and highly built-up areas in the north of Linz, has a limited influence on the remaining sewer system due to the presence of a pressurised collector under the Danube River. This collector is controlled by means of a set of pumps with a maximum capacity of 1.2 m^3s^{-1}. The CSO detention tank (CSODT1 in Figure 25) which is located in this sub-catchment has a storage capacity of 7,500 m^3 for stormwater management. The sub-catchments 2 and 3 are highly urbanised with an industrial area in the East. These sub-catchments share a CSO (CSO 1 in Figure

25) structure regulated with a moveable gate, which discharges into the Danube River. The sub-catchment 4, a combination of residential and industrial areas, has a CSO structure over the Danube as well (CSO2 in Figure 25). Sub-catchment 5 is mainly a residential area with a CSO detention tank (CSODT3 in Figure 25) with a total volume of 9,400 m^3 allowed to spill into a tributary of the Traun River. There is another CSO detention tank before the WWTP Linz-Asten (see CSODT2 in Figure 25) with a storage capacity of 54,000 m^3. Some ecological problems have been observed in the "Mitterwasser" watercourse as a consequence of overflow discharges from the combined sewer overflow detention tank CSODT2. The Linz AG – Wastewater in cooperation with the Federal authority of the Upper Austrian government are working towards preventing these overflow events by means of rule-based real time control (RTC) in the sewer system (Hochedlinger et al. 2006).

As part of this effort, one UV-VIS spectrometer was installed in the main sewer collector of the sub-catchments 2 and 3 directly in the CSO1. These measurements include 5-minute frequency TSS and COD concentration measurements between March and October 2006. Using the COD and TSS concentration data over 38 dry days, the results of the 3rd order Fourier series model again suggest a near-normal distribution of errors (Figure 26a and Figure 26b) and that the serial dependence was well-represented by a 1st order autoregressive model (Figure 26c and Figure 26d).

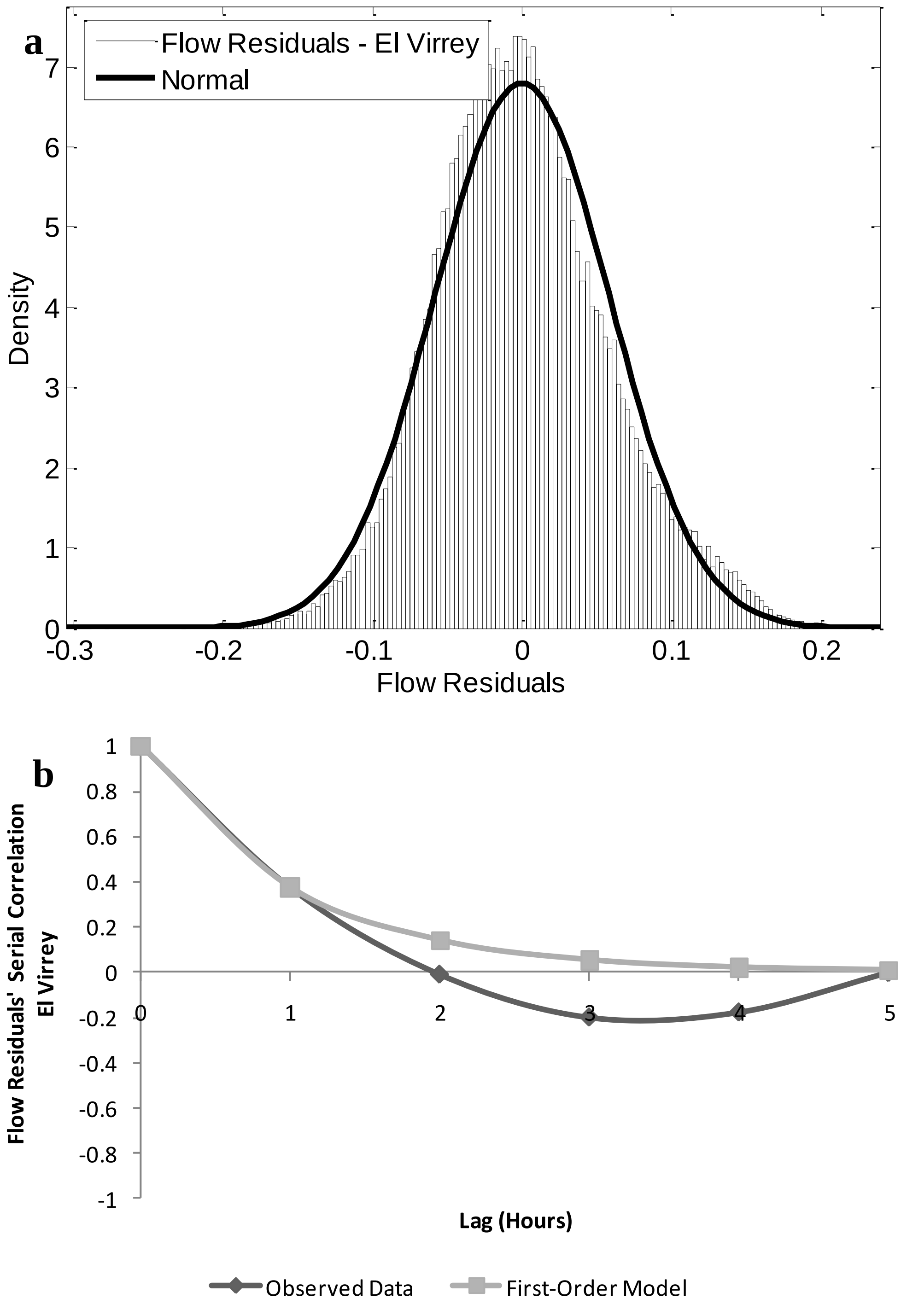


Figure 24. Results for the El Virrey sub-catchment: (a) The histogram of flow errors compared to a fitted normal distribution, (b) Serial correlation coefficients of flow compared to a 1st order linear model

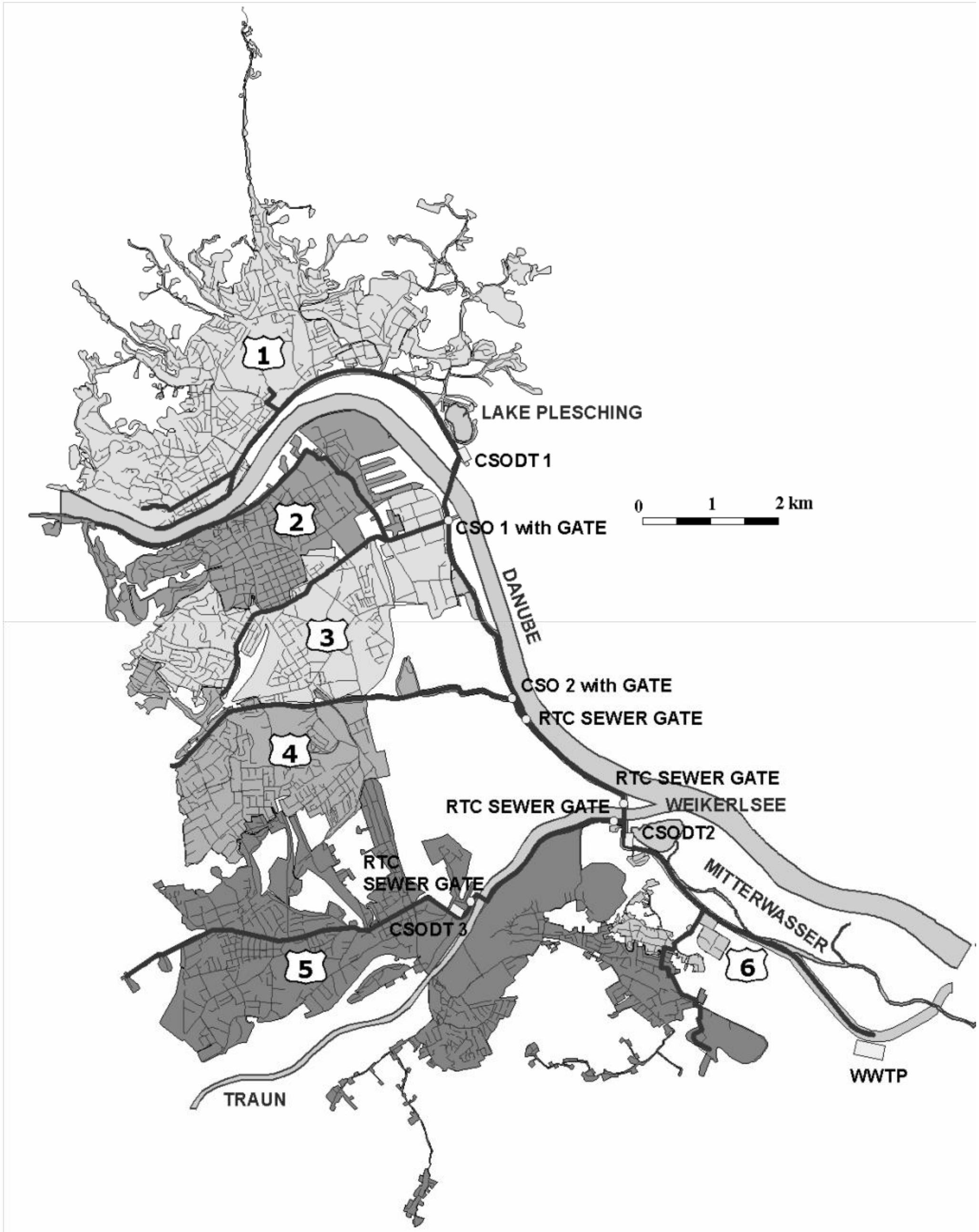


Figure 25. Linz's urban drainage system

Table 7. Main sub-catchments in Linz's sewer system

Sub-catchment	Catchment Area	Sewer
No.	**Km^2**	**Km**
1	12.32	162
2	5.78	69
3	6.89	69
4	5.74	71
5	15.39	162
6	2.03	26
Total	***48.15***	***559***

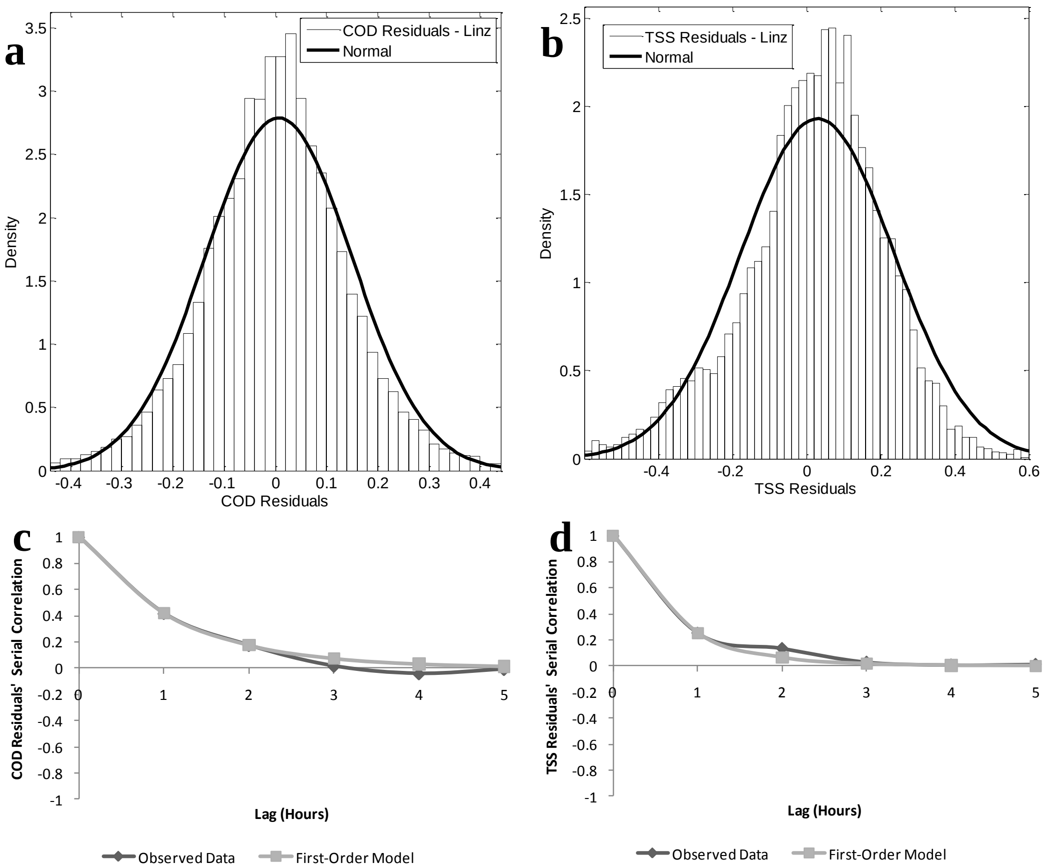


Figure 26. Results for the Linz sewer system (Austria): (a) The histogram of COD errors compared to a fitted normal distribution, (b) The histogram of TSS concentration errors compared to a fitted normal distribution, (c) Serial correlation coefficients of COD compared to a first-order linear model, (d) Serial correlation coefficients of TSS compared to a 1st order linear model

4.7.3 Sub-catchment scale modelling: coupling sub-models a to d

1,000 error realisations (simultaneously for the mean values and the daily variation sub-models) allowed the estimation of distributions of modelled variables for comparison with those observed at the five test sites, A, C, D, F and G. Figure 27 presents results for BOD daily variability (only BOD is presented as similar results were obtained for COD and TSS). It can be concluded that even though the predicted 95 % confidence bounds contain most of observed data, modelled signals exhibit large variability in all sub-catchments. As a consequence, the NSE^{prob} performance indices are rather low (ranging from <0 up to 0.62) (see Figure 27). This is a result of the high variance of the predictions. This may be due to the quasi-random usage of different appliances within different water use sectors (Butler and Graham 1995) or due to the limitations of the model in explaining the patterns, or a combination of these.

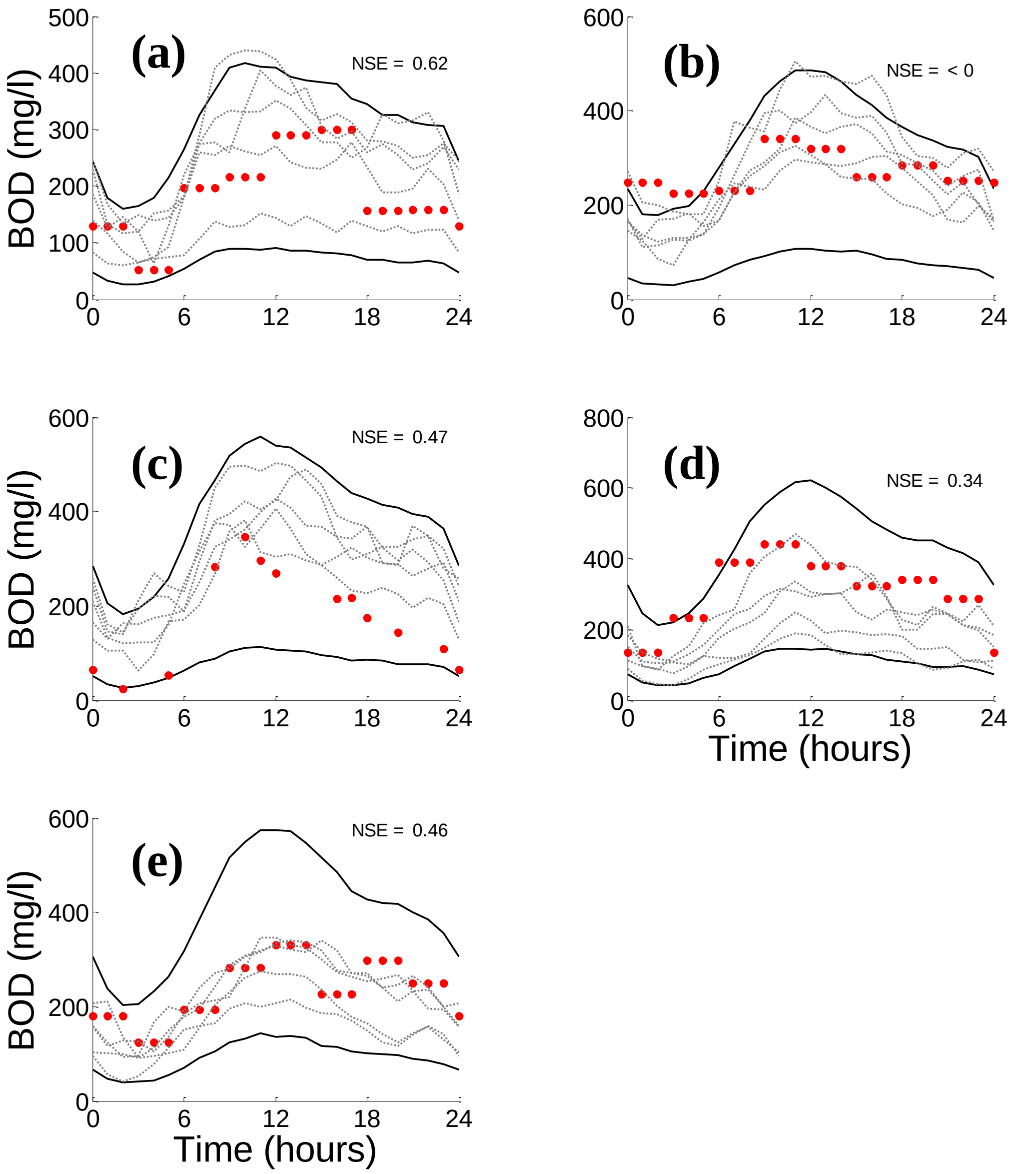


Figure 27. Results for BOD concentration at sites (a) A, (b) C, (c) D, (d) F, (e) G. All plots show observed data (dots), modelled 95% confidence bounds (black lines), 5 examples of model realisations (dotted lines). The NSE reported in this figure is the NSE for probabilistic results proposed by Bulygina et al. (2009))

To complement the assessment of Figure 27, the model's ability to characterise relevant statistics of observed pollutant loads was studied. Figure 28 shows results for standard deviation, maxima and mean values of BOD, COD and TSS pollutant loads, both observed and modelled. These statistics are taken over the 1,000 realisations (i.e. the sample of maxima are the 1,000 daily maxima produced by the 1,000 realisations). Figure 28 shows a relatively good agreement (i.e. most of the observed values lying within the modelled 95% bounds) for all sub-catchments thus indicating that the model is in general capable of representing the observed spatial variability at areas with different sizes, type of sewer system (including wrong connections if the system is separate), number of inhabitants and water use consumption distribution among residential, commercial and industrial uses. However, this conclusion is constrained by the limited available data which do not allow the assessment of modelled variability as monitored data is available over one 24-hour period at each sub-catchment. Further monitoring at sub-catchment outlets over many 24-hour periods could allow such evaluation.

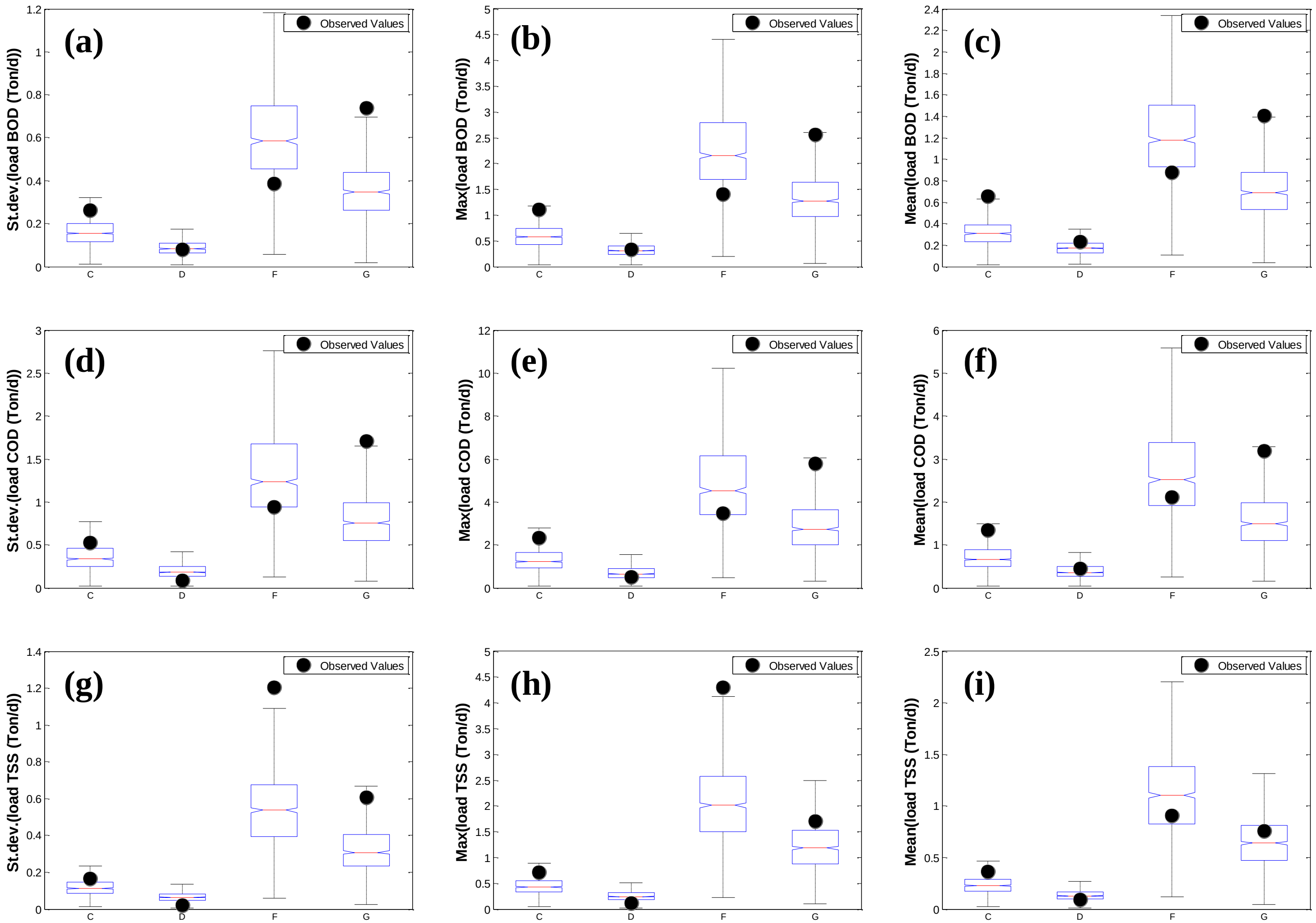


Figure 28. Comparison of observed and modelled BOD, COD and TSS load: standard deviation, maxima, and mean at selected sub-catchment outlets (box plots represent 95% intervals from the 1,000 simulations)

4.8 Summary

The potential for using the proposed mixed deterministic/stochastic generator to characterise gauged and ungauged sub-catchment wastewater outflows is demonstrated for both flows and pollutant concentrations even when supported by low to mid time resolution data sets. However, the proposed methods strongly depend on the quality of available data sets. Despite the fact that the stochastic modules aim to represent the inherent uncertainty in the required input data (e.g. the water user database, the mean and sub-daily variations of flows and pollutant concentrations values at gauged sub-catchments) additional monitoring is needed to further validate the generator. Furthermore, most of the observed diurnal profiles correspond to 3-hour average values. In this Chapter, it was shown that there is limited sensitivity of using hourly or 3-hour average values for residential sub-catchments. However, further data is needed to quantify such sensitivity in cases of more dynamic water uses (i.e. highly industrialised sub-catchments). This chapter has also presented a partial assessment of model transferability to other case studies. However, this evaluation does not cover the transferability of the sub-models a, b and c, or the ungauged site characterisation in sub-model d, for which additional extensive spatial data sets would be required. Besides this, rather simplified approaches were used to characterise the water supply return factor and the infiltration flows into the sewer system. Consequently, these aspects can be further improved if additional data is available for the Bogotá case or other case studies.

CHAPTER 5:
MODELLING TO SUPPORT WASTEWATER SYSTEM MANAGEMENT

5.1. Introduction

Urban drainage system models can be useful to assess and manage system performance and to plan its development. However, due to data and computational costs, sophisticated, high-resolution contemporary models of the sewer system may not be applicable. This constraint is particularly marked in developing country mega-cities where catchments can be large, data tend to be scarce, and there are many unknowns, for example regarding sources, losses and wrong connections. This Chapter presents work to develop a suitable modelling framework to support operation and development of the wastewater system of Bogotá (Colombia) under dry weather conditions. Specifically, it is pursued the identification of parsimonious models for application to large and complex sewer systems for the purpose of providing inputs to WWTP operation by means of studying a variety of alternatives to represent system's spatial and temporal variability including in-sewer processes.

5.2. Sewer system model

(Achleitner et al. 2007) developed the City Drain toolbox for urban drainage system modelling. The sewer system model is built within the toolbox using conceptual blocks representing the source, the sub-catchment (both combined and separate systems), the sewer, the wastewater treatment plant and the river. The sewer blocks allow for flow and pollutant routing using Muskingum routing, although one of the main advantages City Drain is the possibility of modifying the code according to specific needs. City Drain has been previously applied for different purposes and to different case studies (Achleitner et al. 2005; Achleitner et al. 2007; Achleitner and Rauch 2007; De Toffol et al. 2007; Engelhard et al. 2008; Achleitner et al. 2009; Kleidorfer et al. 2009; Rodriguez et al. 2010; Manz et al. 2012).

A version of City Drain is used for the Bogotá case study to model the main sewer system, from sub-catchment outlets to the inlet of the Salitre WWTP. In order to facilitate model development, a schematic description of the urban drainage, including components and interactions, was performed using a GIS platform (Figure 29). The resulting model of Salitre includes 180 sub-catchments drained by a number of main sewer pipes, 31 CSOs and 4 wastewater pumping stations. Physical properties and operational data of CSOs and pumping stations were provided by the EAAB. The sub-catchment is modelled using an hourly time step which considered sufficient to capture DWF dynamics.

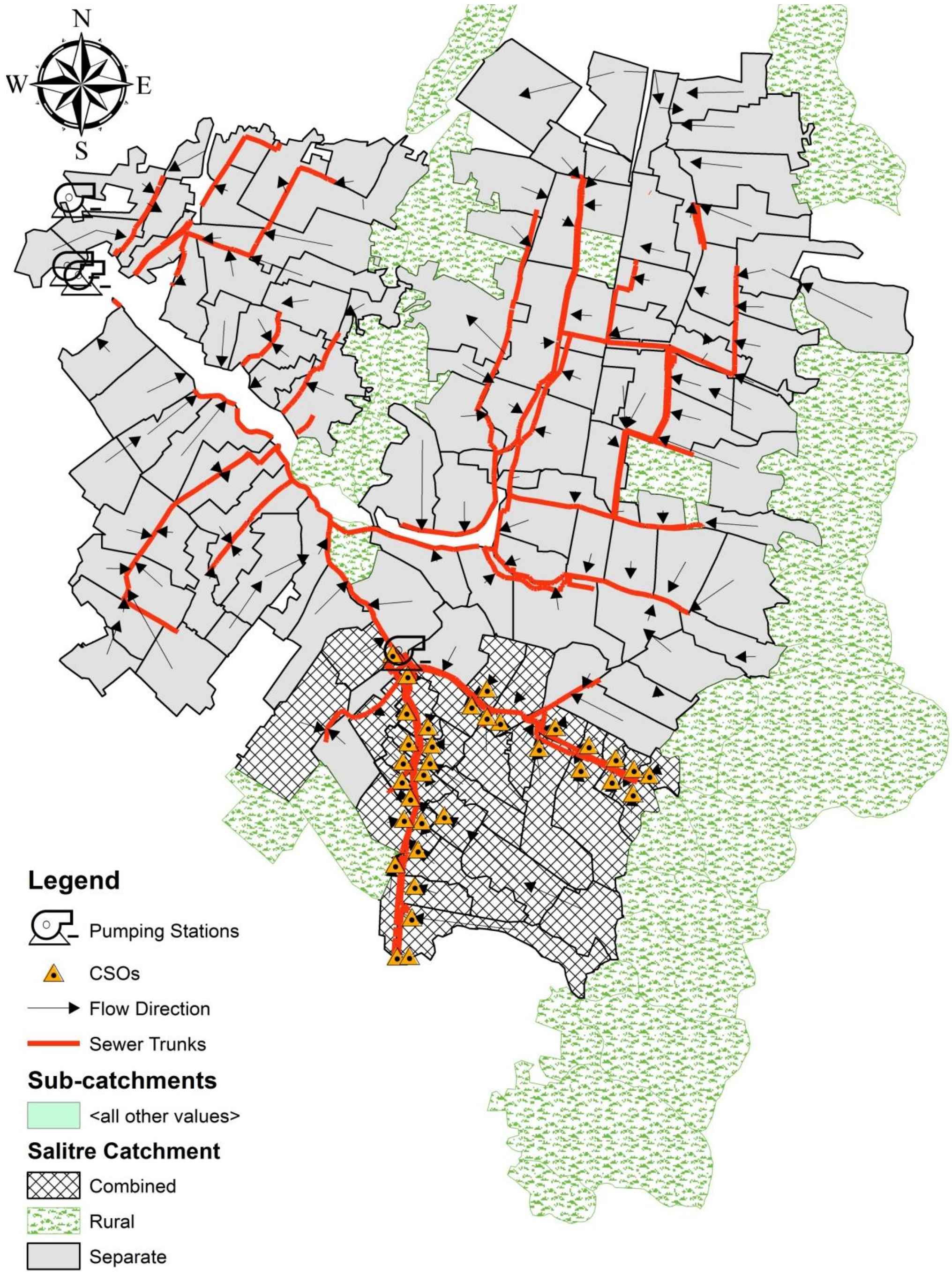


Figure 29. Schematic representation of the Salitre catchment

5.3. Integration and evaluation

5.3.1 Monte Carlo analysis

The three components described above – the database (see Chapter 3), the pollution load model (see Chapter 4) and the network model - allow the application of the City Drain toolbox to dynamically model the Salitre catchment. As it is considered beneficial to represent wastewater DWF patterns using both deterministic and stochastic components (the benefits in doing so are demonstrated later in the Chapter), multiple realisations of all stochastic components are needed in order to characterize the expected system variability. Each realisation represents a possible time series according to the model. If the model is adequate, in both its deterministic and stochastic components, the observed data will appear to be one realisation. In this work a set of 1,000 error realisations was considered sufficient to characterise error variability. Although the model may be used to generate longer time series, each realisation here consists of a 2-day run: the first day to initialise the model, and the second to assess the model against observed diurnal patterns.

5.3.2 Validation

The validation analysis aims to quantify the performance of the semi-distributed model when replicating observed flows and pollutant concentrations within the main sewer (both combined and separate) trunk system. 8 gauged sites (out of the 13 sites described in Table 3) aim to support the proposed assessment: sites B and E are located in the combined sewer trunk system, sites H and I corresponds to CSO structures (remind that there are permanent CSO discharges into stormwater open channels during dry weather), sites J, K, L and M characterises the sanitary trunk system in the separated part.

5.3.3 Sensitivity analysis

It was reviewed in Chapter 2 that in many sewer systems, hydraulic retention provides sufficient time for the wastewater and solids to be transformed by in-sewer processes. Besides this, it was stated that the there are three variables that largely control the rate and nature of pollutant transformations: (a) presence of DO, (b) bulk wastewater temperature and (c) hydraulic retention time. Despite the conclusion made in Chapter 3 which indicates that for the Bogotá sewer system the reaeration process is unlikely to be a main driver of in-sewer transformations and that due to relatively high temperatures anaerobic conditions are expected to prevail within the system, this Chapter aims to quantify the need of representing in-sewer transformation processes within the Salitre Catchment. For this, in-sewer pollutant reduction rates reported by (Almeida et al. 2000) for BOD, COD and TSS are used in this work. Even though their rates correspond to aerobic conditions (due to a relatively high average slope of 0.7%) and thus more pronounced pollutant reductions (in comparison to the expected anaerobic conditions within the Salitre catchment), they are considered useful in the context of this work in order to assess model outputs sensitivity when transformations are included. Two additional sets of 1,000 model iterations were carried out, one using the reported average reductions (i.e. 4% h^{-1}, 6% h^{-1}, -6% h^{-1} for BOD, COD and TSS respectively) and other including a random sampling of the rates according to their reported standard deviations (i.e. 10% h^{-1}, 7% h^{-1}, 19% h^{-1} for BOD, COD and TSS respectively). These two sets are named here Model B and Model C respectively. Results are compared against those from the model without inclusion of the original pollutant degradation (Model A). It is not only important to identify which model alternative (from A, B or C) is justified regarding in-sewer processes. A lumped (i.e. spatially averaged) model (Model D) without in-sewer processes was also tested in order to examine the value of the semi-distributed model.

5.4. Results

5.4.1 Model A: semi-distributed model with no in-sewer transformations

8 sites within the main sewer system are used here to contrast model results against available observations: sites B, E, H, I, J and K (see Table 3). Results from 1,000 model iterations are in Figure 30 and Figure 31. Figure 30 presents results for TSS sub-daily variability (only TSS is presented as similar results were obtained for BOD and COD). Again, there is considerable uncertainty in modelled DWF patterns and low NSE^{prob} values (values ranging from < 0 to 0.73). It is also noticed that the larger the hydraulic retention time the less pronounced the predicted variability, likely due to hydraulic attenuation.

There is little evidence from Figure 30 that this simple water quality model (recall that the sewer model used here does not account for in-sewer processes) cannot represent diurnal profiles in the combined system including CSOs (sites B, E, H and I, see Figure 30a to Figure 30d) and in the separate system (sites J and L, see Figure 30e and Figure 30g). Most of observed data lie within the simulated 95% intervals at varying retention times (up to 5.6 hours as the case if site L, see Figure 30g). However, there is evidence of disagreement when looking at sites K and M (Figure 30f and Figure 30h). These two sites correspond to main sewer trunk outlets to a 600 m open channel connecting to the Salitre WWTP. The open channel provides additional retention capacity depending on the WWTP pumping operation. Recall that the Salitre WWTP has a capacity of 4 m^3s^{-1}, and the dry weather flows exceed WWTP treatment capacity during about 60% of the time thus triggering backwater phenomena. This causes increased retention time (over the expected 6.7 hours) and increased significance of in-sewer processes.

Additional analyses are in Figure 31, showing results for standard deviation, maximum and mean values of BOD, COD and TSS pollutant loads, both observed and modelled. Again there is good agreement for mean values. However, there are some discrepancies for standard deviations and maximum values particularly regarding TSS. These results suggest that resuspension/sedimentation processes are more relevant than those transformation processes that alter BOD and COD concentrations.

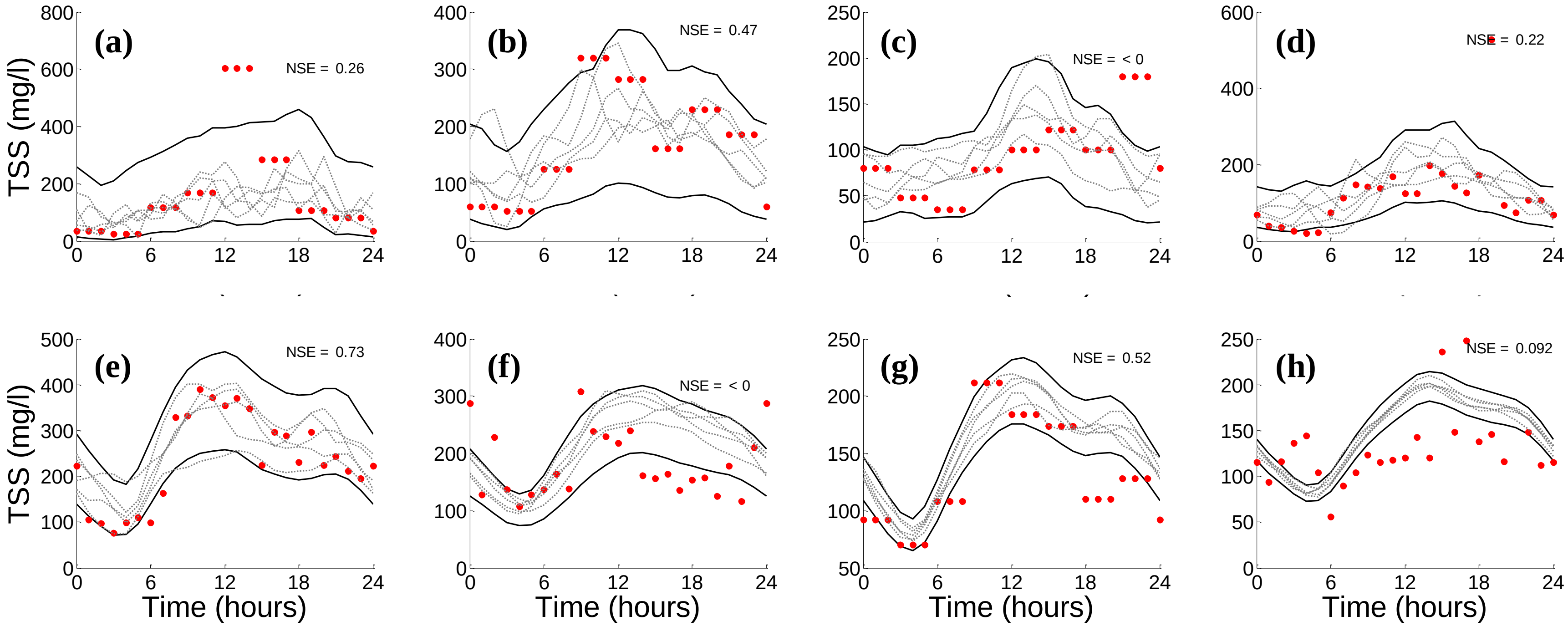


Figure 30. Results for TSS concentration at sites (a) B, (b) E, (c) H, (d) I, (e) J, (f) K, (g) L and (h) M. All plots show observed data (dots), modelled 95% confidence bounds (black lines), 5 examples of model realisations (dotted lines). The NSE reported in this figure is the NSE for probabilistic proposed by (Bulygina et al. 2009))

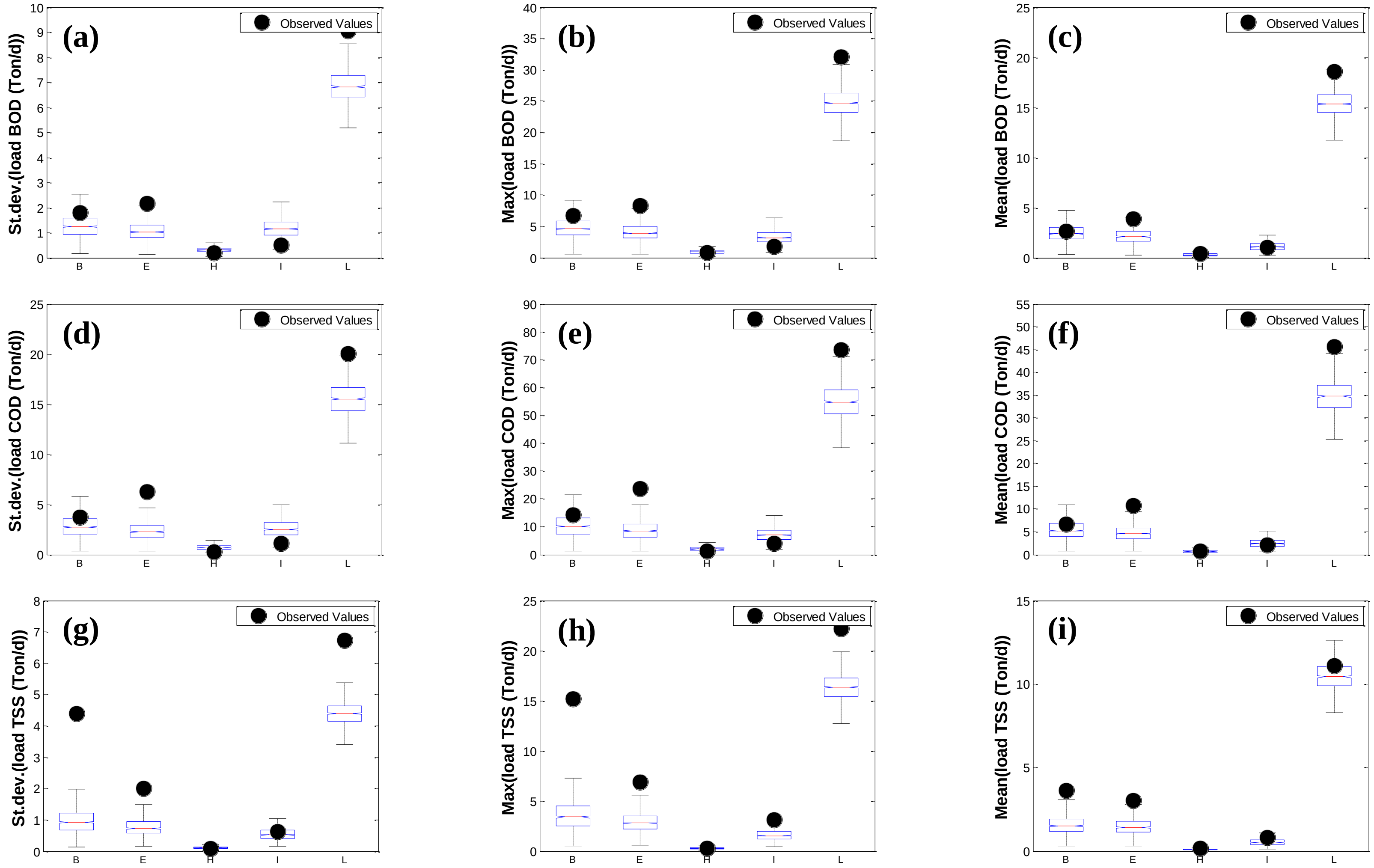


Figure 31. Comparison of observed and modelled BOD, COD and TSS load: standard deviations, maximums, means at selected sites within the main sewer system (box plots represent 95% intervals from the 1,000 simulations)

The importance of in-sewer processes under some circumstances are also seen when comparing model results with the EAAB monitoring data from sites E, L and M. Figure 32 shows that the model provides a good representation of observed concentrations at sites E and L, however at site M, when treatment capacity is exceeded (around 10 am to 10 pm), in-sewer transformation and sedimentation processes seem to become relevant; while, when capacity is not a limiting factor, high concentrations are observed at site M, most likely due to re-suspension of deposited material.

5.4.2 Models B to D: sensitivity to altered model complexity

Figure 32 also present results from Models B and C (i.e. Model B: semi-distributed model where average transformation rates are implemented, Model C: semi-distributed model where transformation rates uncertainty is included). It can be concluded that even for relatively large hydraulic retention times (i.e. 5.6 hours, Figure 32d to Figure 32f), neither inclusion of average transformation rates nor transformation variability seems to greatly improve model performance. This is because uncertainty in DWF profiles within the sewer system is largely controlled by pollutant load variability at the sub-catchment scale; while for site M, severely affected by backwater effects, the processes are more complex than considered in any of the models tested here.

Model D corresponds to a spatially averaged model without in-sewer processes representation. Model D can be considered as the simplest sewer system model in the context of this work. Results in Figure 32 show that the spatially averaged model tends to overestimate pollutant concentrations between 6 am and 12 pm. However, there is not marked differences from 12 pm if contrasted to the semi-distributed alternatives. This result implies that in cases where there is not detailed information about the main sewer system properties, a lumped model can still provide valuable information of approximate diurnal dynamics of pollutant concentrations. Nevertheless, an over estimation of concentrations and loads is expected. In conclusion, except for the special case of site M, the simplest semi-distributed model (Model A) is considered to be suitable for modelling the network.

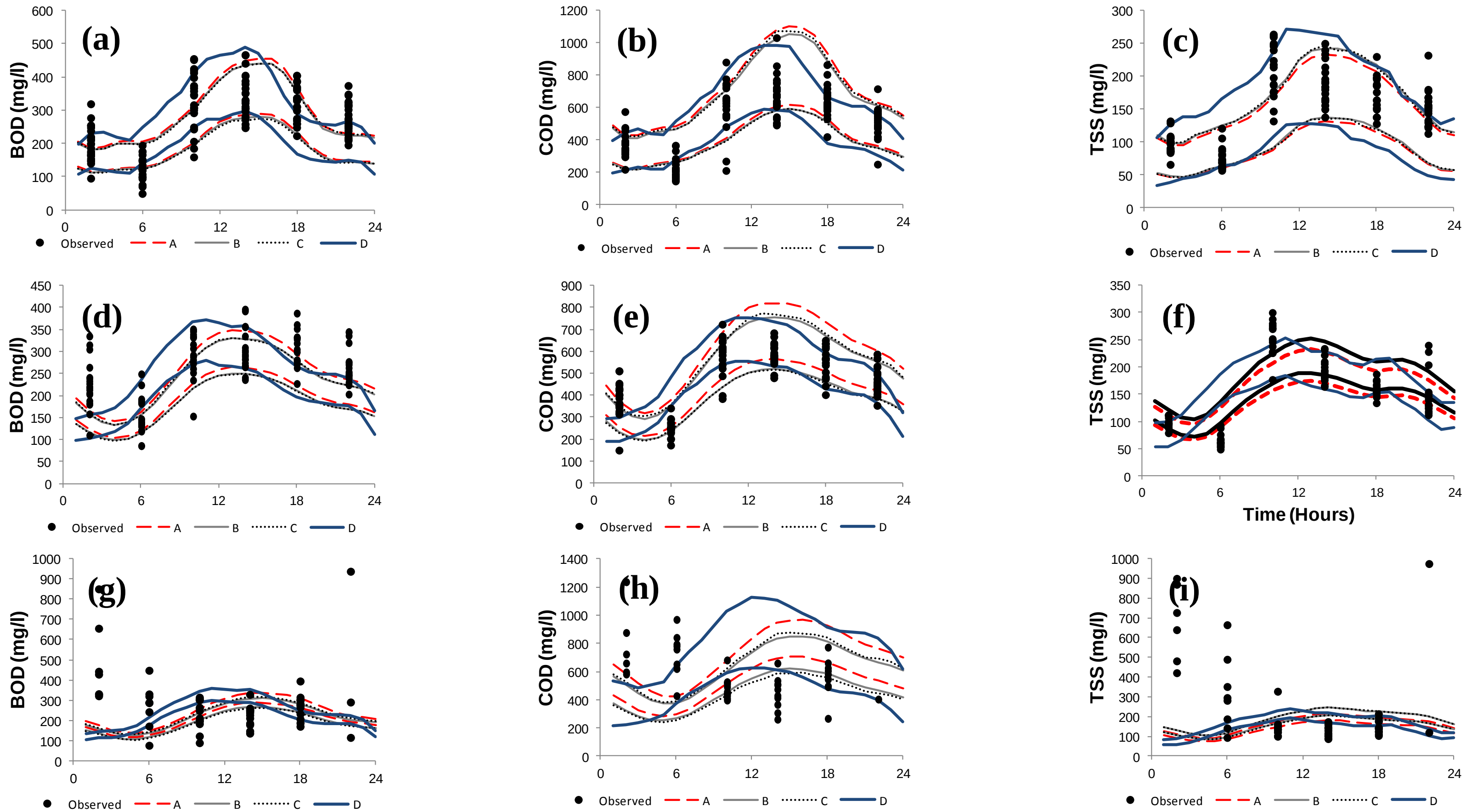


Figure 32. Observed BOD, COD and TSS concentrations and modelled 95% confidence bounds for: (A) semi-distributed model with no in-sewer transformations, (B) semi-distributed model with average transformation rates, (C) semi-distributed model with random transformation rates and (D) spatially averaged model without in-sewer process representation. Figures (a) to (c) correspond to downstream site E, while (d) to (f) to downstream site L, and (g) to (i) to site M respectively

5.5. Summary

Results from a catchment within Bogotá, area 150 km^2 and with 2.5 million inhabitants, show that the model outputs capture the scale and dynamics of the observed concentrations and loads at various points in the sewer system. However uncertainty is high because much of variability of observed dry weather flow profiles is apparently random. Against this variability, the effects of in-sewer processes were not identifiable except where backwaters caused particularly high retention times. Hence the work has resulted in an operational model with a scientifically justified, yet useful, level of complexity for Bogotá. More generally, the work demonstrates the insights into suitable level of model complexity that may be gained by uncertainty and sensitivity analysis.

CHAPTER 6:

SEWER SYSTEM MAINTENANCE

6.1 Introduction

The principal objectives of the Chapter are: (a) to understand which physical properties of the sewerage system are affecting blockage rates, (b) within the constraints of the available data; to seek evidence of trends in blockage rates; and (c) to identify statistical models (i.e. homogeneous and non-homogeneous Poisson process models) that provide the most appropriate descriptions of the random nature of blockages within each grid square. To do so, this Chapter analyses an exceptionally long (7.5 years) and spatially detailed (9,658 grid squares – 0.03 km^2 each – covering a population of nearly 7.5 million) sediment-related blockage data set obtained from a customer complaints database in Bogotá (Colombia). In this Chapter, "effective blockage" refers to effective sediment-related failures and blockage rate refers to the average number of effective blockages per year as calculated in each square.

6.2 The models

This Chapter describes a set of statistical models of sediment-related failures based on the Ascher and Hansen (1998) framework. The sewer system is assumed to be a repairable system where data of failure inter-arrival times can be derived from the times of reported blockages. This section also briefly describes different data analysis methods that were applied in this research aiming to characterise the observed spatial variability of blockage rates.

6.2.1 Homogenous Poisson process and renewal process models

A natural first hypothesis is that sediment-related blockage events occur completely randomly in time, which can be represented by means of a homogenous Poisson process (HPP). The HPP builds upon a fundamental assumption of complete lack of memory: a failure is equally likely to occur at any time regardless of the physical state of the system and the system's history of failures. This implies that the system is neither improving nor wearing out with age, but rather is maintaining a constant intensity of failure (Crow 1975). Let λ be a constant which measures the mean rate of occurrence of failure events. Consider the number of events, N_t, occurring in an arbitrary time interval of length t. Then N_t has a Poisson distribution of mean λt,

$$\mathrm{prob}(N_t = r) = \frac{(\lambda t)^r e^{-\lambda t}}{r!} \ (r = 0, 1, \dots) \qquad \text{Equation 14}$$

If T' is the interval from a time origin to the next event the cumulative distribution function (cdf) and the probability density function (pdf) of T' are (Cox and Lewis 1966),

$$F_{T'}(T') = 1 - e^{-\lambda T'} \ (T' \geq 0) \qquad \text{Equation 15}$$

$f_{T'}(T') = \lambda e^{-\lambda T'} \ (T' \geq 0)$ Equation 16

The origin from which T' is measured may be any random point in time since the previous failure, although for convenience is defined here as the time of the previous failure. It follows that T1, T2, etc. are the intervals between events and are mutually independent random variables described by pdf Equation 16 which corresponds to an exponential distribution.

When the information in the chronological order of inter-arrival times is ignored, they often appear exponential leading to the impression that the system can be modelled using a HPP whereas in fact there is non-stationarity. This can be avoided by applying an appropriate test for trend before attempting to fit a distribution to the inter-arrival times. The Laplace test has been recommended as a suitable test to determine whether an HPP is a justified model (Cox and Lewis 1966; Gaudoin 1992; Ascher and Hansen 1998). A Laplace test statistic bigger than 1.96 signifies that system reliability is deteriorating and if smaller than -1.96 signifies that reliability is improving, in each case at the 95% confidence level, implying that an HPP model is unlikely to apply. If no evidence of trend is concluded, a goodness-of-fit test (e.g. the total time on test (TTT) statistic, Epstein and Sobel (1953)) can be used to determine if the HPP model is an appropriate description of the inter-arrival times. If not, a more general RP model can be tested, e.g., using a Weibull or gamma distribution (Ascher and Hansen 1998; Korving et al. 2006). For example, it might be supposed that the inter-arrival times of sediment-related blockage events are described well by a two-parameter Weibull or gamma RP model if it is expected that the blockage rate is low in the period immediately after cleaning and then increases over time, and this process is reset after cleaning. On the other hand, a non-stationary model is likely to be better in cases where the sewer becomes more prone to blockages from one blockage event to the next, for example due to structural deterioration or increasing sediment load.

6.2.2 Non-homogenous Poisson process models

A generalization of the HPP which allows for changes or trend in the intensity of system failures is the non-homogenous Poisson process (NHPP) (Crow 1975), where the cumulative number of failures is assumed to follow a Poisson distribution, but the mean number of failures is not directly proportional to time t. Crow (1975) proposed a NHPP model in which the intensity function is a power law – known as Crow's model – represented by,

$\lambda = \beta_0\beta_1 t^{\beta_1-1}$, $\beta_0 > 0$ and $\beta_1 > 0$ Equation 17

where λ is the time-dependent mean failure rate; β_0 is a scale parameter; β_1 is a growth parameter. For $\beta_1 > 1$ the failure rate increases over time. For $\beta_1 < 1$ the failure rate decreases and the system improves.

Alternatively, Cox and Lewis (1966) proposed a log-linear model – known as the Cox and Lewis' model – which was originally developed to describe improving systems (Korving et al. 2006) and is described by,

$\lambda = \exp(\beta_0 + \beta_1 t)$ Equation 18

where β_0 is the scale parameter, β_1 is the growth parameter and the failure rate increases over time for $\beta_1 > 0$ and decreases for $\beta_1 < 0$. The parameters of both Equation 17 and Equation 18 can be estimated by maximum likelihood (Korving et al. 2006). The growth models of Crow and Cox and Lewis are neither suitable for describing rapid changes in the failure rate, for example where failures are clustered, nor for periodically changing failure rates. As stated by Korving et al. (2006), they are restrictive due to their monotonic failure rate (time increasing or decreasing). Although the implementation of a more complex model for example using a stochastic mean failure rate (e.g. doubly stochastic Poisson process) may be appealing, the increased number of parameters may lead to unacceptable parameter uncertainty (Korving et al. 2006).

6.3 Spatial analysis of blockage rates

Despite (*a*) the highly complex physical mechanisms of sediment deposition; (*b*) the number of factors that may contribute to a sediment-related blockage and (*c*) the disagreement on which physical properties can be considered as influential factors, previous research has identified simple models for mean blockage rate estimation. In this work, some of those physical properties that have been previously identified as influential - such as pipe diameter, slope and length - are available and used in the analysis described below. Each property was individually studied for model identification together with a factor that simultaneously combines them (see Equation 19) as proposed by (Savic et al. 2005) and named here the COST-S factor:

$$\mathrm{COST-S\ factor} = \mathrm{Length} \cdot \mathrm{Slope}^{1/2} \cdot \mathrm{Diameter}^{2/3} \qquad \text{Equation 19}$$

These models can potentially be used to explain the spatial variability of blockage rate and to enable estimation of rates for unobserved parts of the system (as reviewed above). In order to identify (if any) explanatory models of the observed blockage rates, different data analysis techniques were used in this work:

- **Stepwise regression analysis and the multi-objective search technique EPR MOGA (Giustolisi and Savic 2006; Giustolisi and Savic 2009)**: described in Chapter 4.

- **Fuzzy clustering:** clustering is the process of dividing data into classes (e.g. areas of a particular sewer system with similar mean pipe slope, pipe diameter and total pipe length) or clusters so that elements in the same cluster are as similar as possible, and in different clusters are as dissimilar as possible (Mirkin 2005). In fuzzy clustering each element has a degree of belonging to all clusters rather than belonging to just one. The Matlab built-in fuzzy c-means procedure (MathWorks) is used here in order to identify appropriate data subsets to which further apply the stepwise regression and EPR MOGA analysis techniques.

6.4 Results

6.4.1 Relation between the sewer system properties and the rate of blockages

Non reported vs. reported squares

Out of the total of 9,685 squares, 929 – with reported sewer infrastructure – do not have any reported blockages over the 7.5 year period, representing approximately 10% of the total number of squares. To analyse if sewer system properties in these "problem-free" squares differ from those in the other squares, Figure 33a and Figure 33b present the two sets of probability density functions (pdf) for mean pipe diameter in problem-free squares and in all other squares where blockages were recorded, respectively. Figure 33c and Figure 33d show the same comparison for mean slope (in m/m), while Figure 33e and Figure 33f show results for total pipe length. To interpret Figure 33, the two-sample Kolmogorov-Smirnov test (Massey 1951) was applied to assess if there are or there are not significant differences between the statistics of the sewer system properties in non-reported and reported squares. At the 95% significance level, results indicate that diameter, slope and total pipe length have a significant influence on the propensity to blockage.

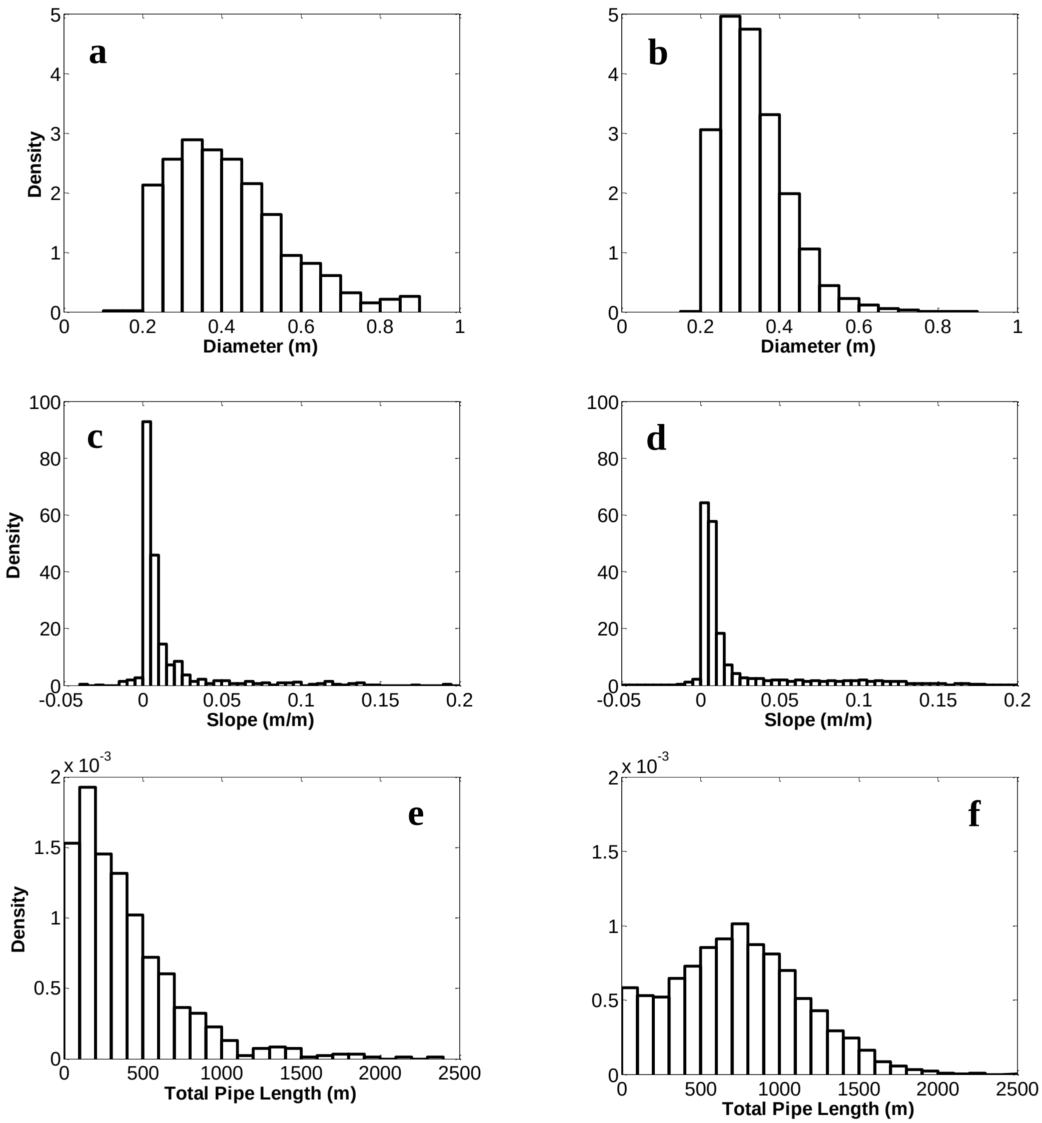


Figure 33. (a) Mean pipe diameter pdf for non-reported squares; (b) Mean pipe diameter pdf for reported squares; (c) Mean pipe slope pdf for non-reported squares; (d) Mean pipe slope pdf for reported squares; (e) Total pipe length pdf for non-reported squares; (d) Total pipe length pdf for reported squares

Assessing spatial patterns

Based on available historical records, Figure 34a-Figure 34c present spatial data of blockage rates (per structural unit in this case) observed in manholes, gully pots and pipes respectively (again, the pipe structural unit is a 100 m length of pipe). These blockage rates were obtained from effective sediment related customer complaints in only those squares with reported infrastructure. In addition to rates data, Figure 35 presents spatial information of mean diameter, total pipe length, and mean slope for each square in which the blockage rate is larger than zero, along with terrain elevation data (Figure 35a-Figure 35d respectively). It can be concluded that while low and high blockage rates are observed in squares covering a wide range of system properties, there are some spatial patterns that can be subjectively interpreted. High slopes on the periphery of the catchment tend to increase gully pot blockages, while total pipe length tends more to affect pipe and manhole blockage rates. The influence of pipe diameter is more difficult to see from these plots because it is so widely variable across the catchment. These results encourage a more detailed and formal analysis of potential relevant factors, which is presented in the following subsection.

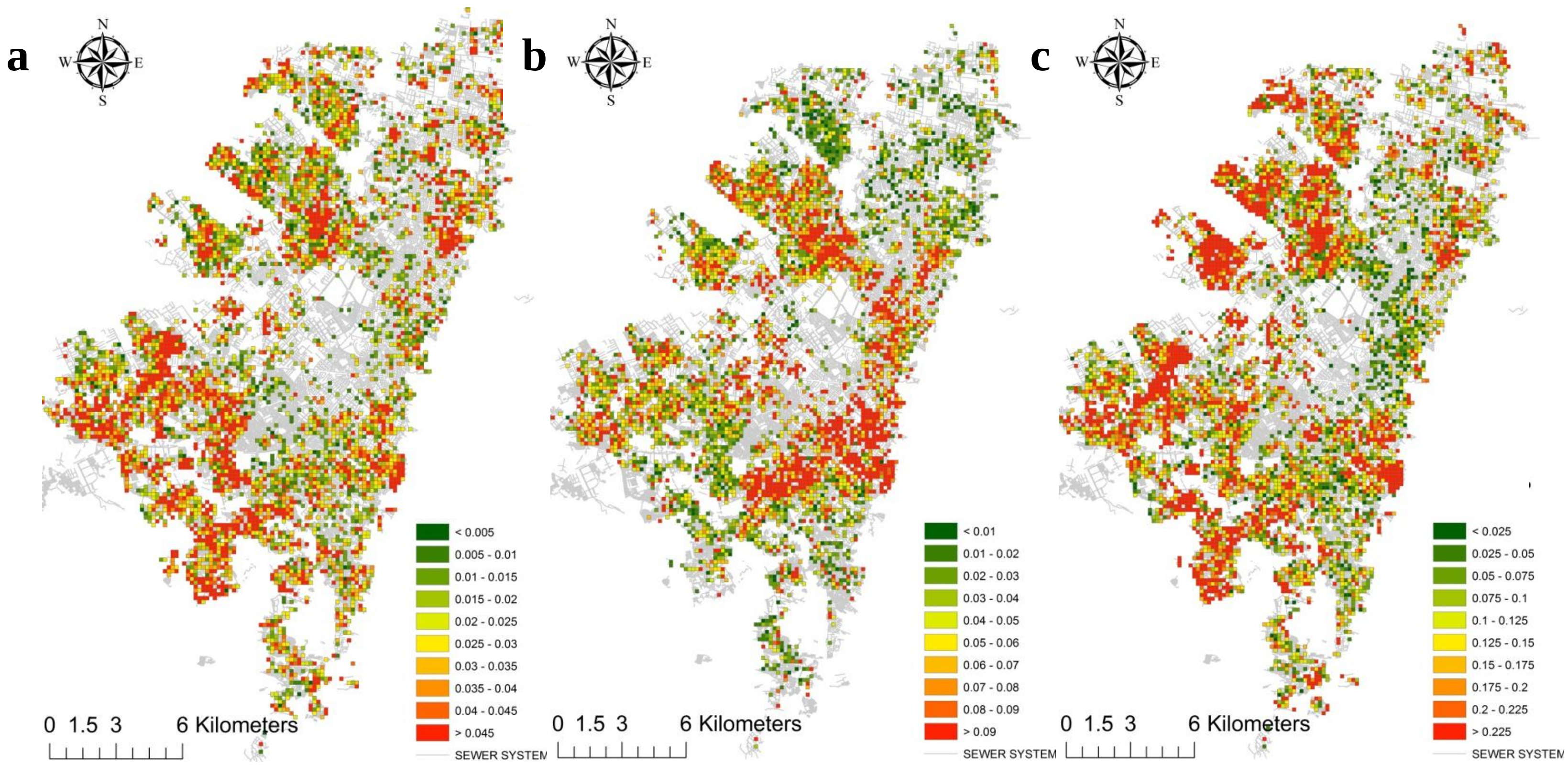


Figure 34. Spatial distribution of blockages rates (per year per structural unit) for (a) manholes, (b) gully pots, (c) pipes

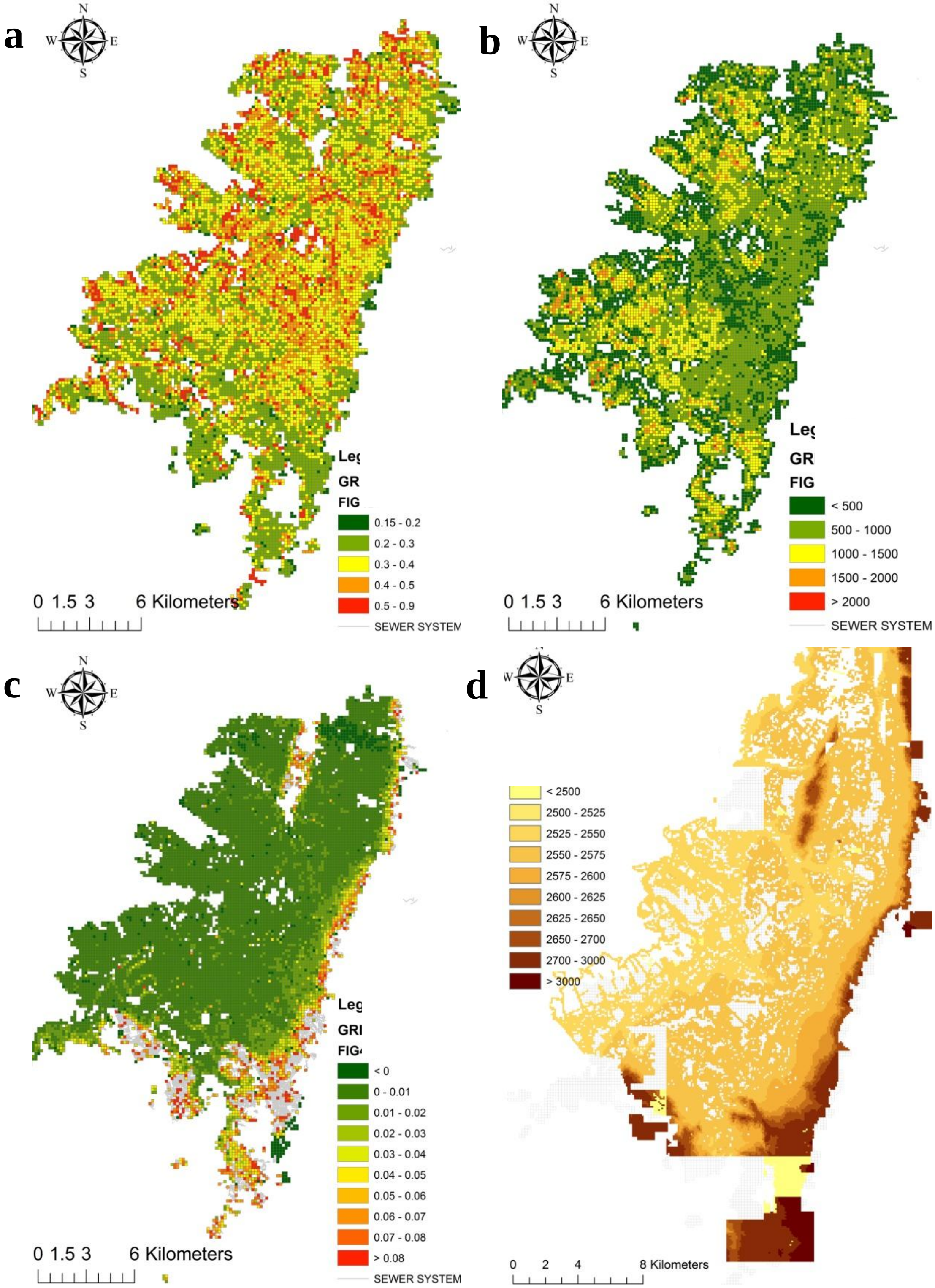


Figure 35. Spatial distribution of system properties for (a) mean pipe diameter (m), (b) total pipe length (m), (c) mean pipe slope (m/m) and (d) terrain elevation (m.a.s.l.)

Factors affecting blockage rates

From Figure 36, it can be concluded that some squares with average pipe diameter between 0.2 and 0.4 m suffer the highest blockage rates not only for pipes but also for manholes and gully pots. The maximum recorded blockage rates tend to decrease as the diameters increase. Most of squares have a mean diameter between 0.2 and 0.4 m (see Figure 33b), so this result might be due merely to better sampling of squares with higher blockage rates. However, a two-sample Kolmogorov-Smirnov test comparing the blockage rate distributions over different ranges of pipe diameters supports the conclusion that squares with average pipe diameter between 0.2 and 0.4 m suffer the highest manhole and pipe blockage rates. Mild slopes have also been commonly associated with higher blockage rates, also observed here. Figure 36 clearly shows that squares with values of average pipe slopes (in m/m) ranging between 10^{-3} and 10^{-2} correspond to squares with high blockage rates. Although it is expected that slopes < 10^{-3} would provide the best conditions for sediment blockages, in this case these slopes are in the lower squares of the catchment and correspond to larger pipe diameters, lowering propensity to blockage. When analysing the relationship between blockage rates and the total pipe length in a square, it seems that the longer the pipe length, the larger the blockage rate.

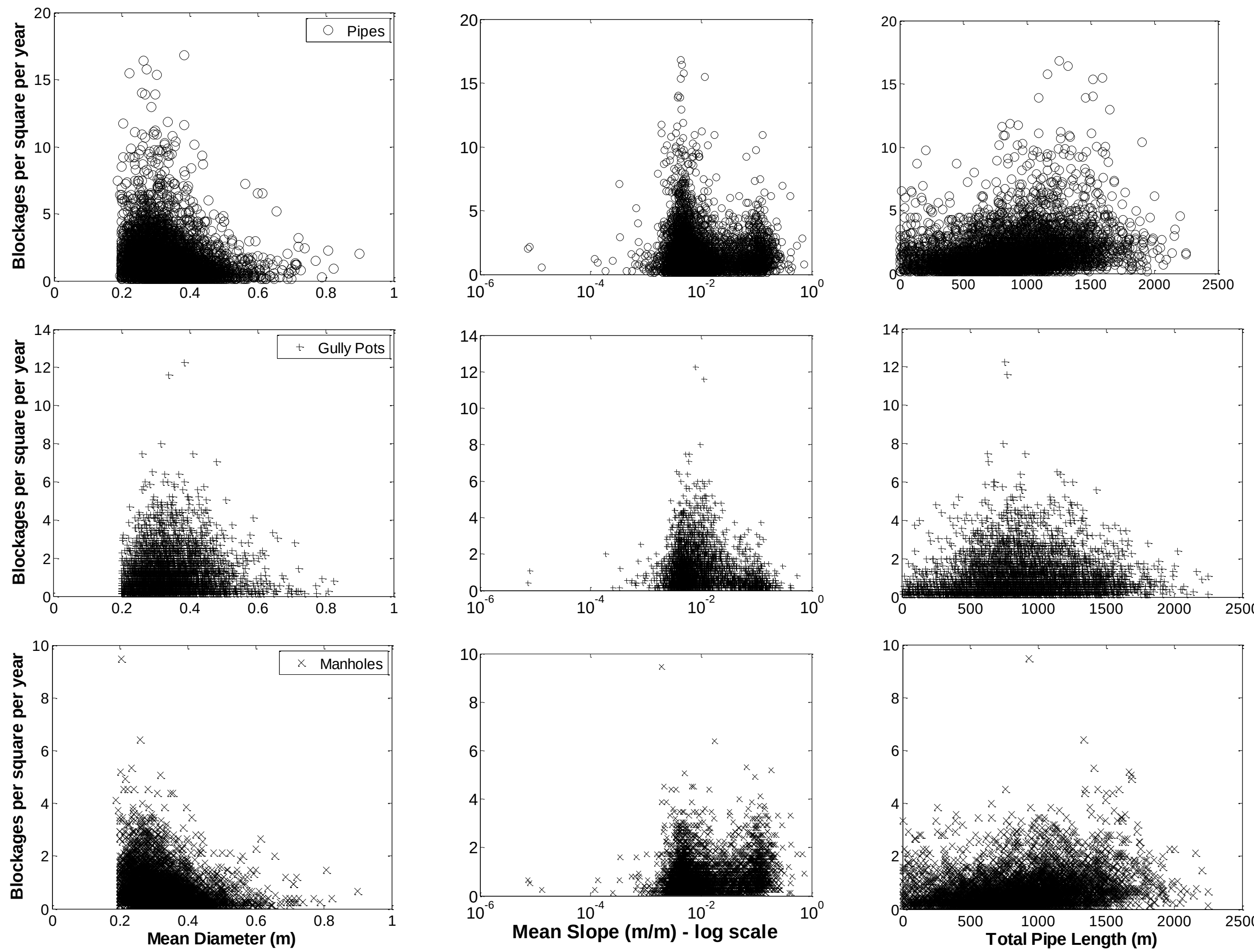


Figure 36. Relationships between sewer system physical properties and observed blockage rates (rows correspond to pipes, gully pots and manholes; while columns correspond to mean diameter, mean slope and total pipe length respectively). Note that the slope scale is base 10 logarithmic

A stepwise regression analysis (using a Box-Cox transformation of both predictors and predictands to improve normality before applying the regression) was carried out in order to identify a relationship between system properties (i.e. diameter, slope, pipe length and number of structural units) and the blockage rate. This was individually performed for pipes (4,816 samples), gully pots (4,287 samples) and manholes (4,472 samples) respectively. Units with zero effective blockages were omitted. It was found that manhole and pipe blockage rates are partly explained by mean pipe diameter, mean pipe slope, total pipe length, number of manholes and the COST-S factor. In contrasts gully pot blockage rates were not related to the slope. However, the obtained multivariate linear regressions provide a poor explanation of the observed annual blockage rates with R^2 of 0.15, 0.06 and 0.16 for manholes, gully pots and pipes respectively.

As linear models are incapable of characterising blockage rates (with the proposed predictands), the EPR MOGA-XL software (Giustolisi and Savic 2006; Giustolisi and Savic 2009) was used to explore if nonlinear models perform better instead. However, mathematical expressions which gave significant R^2 values could not be identified even when clustering the data using the Matlab built-in fuzzy c-means procedure (MathWorks). These results illustrate the complexity and randomness in sewer blockages, and that more explanatory variables should ideally be taken into account to describe observed blockage rates. Further analysis, for example, could include dynamic variables such as structural conditions, high resolution spatially distributed rainfall data and water consumption rates. However, sufficient dynamic data are not presently available.

6.4.2 Tests for linear trends in failure rate

From this point in the Chapter, attention is focused on those squares that have more than 7 blockages over the 7.5 year period. The main reason for this is that in order to perform meaningful statistical tests a minimum number of samples have to be provided (e.g. for applying the Anderson-Darling test used later in the Chapter, a minimum of 8 samples is required). It is important to recognize, referring to Figure 11b, that data availability imposes limitations on how effectively the blockages can be disaggregated into types of structural unit (i.e. pipes, manholes and gully pots). If blockages for all types of units are aggregated then it can be seen that approximately 90% of the squares have more than 7 reported blockages; however, if the aim is to analyse individual structures the total amount of squares with more than 7 blockages is drastically reduced (only about 50%, 30% and 20% of the squares have more than 7 reported blockages individually for pipes, gully pots and manholes respectively).

The Laplace test (Cox and Lewis 1966; Gaudoin 1992; Ascher and Hansen 1998) was applied to quantify if there is statistical evidence of deterioration or improvement in the blockage rate. The Laplace test corresponds to a Poisson process null hypothesis. To assess the proportion of squares for which there is a trend, the cumulative probability of the Laplace test score (U_L) can be found in Figure 37, from which it is possible to conclude that about 69% of the analysed squares hold no trend at the 95% confidence level (-1.96 < UL < 1.96) when blockages in all types of units are considered together. This percentage is 82%, 70% and 66% respectively if manholes, pipes or gully pots are considered separately. According to these results, in most of the analysed squares the inter-arrival time between blockages can be represented by means of probability distributions (e.g. exponential, Weibull or gamma distributions) with stationary parameters. However, there is still a considerable fraction of squares for which non-stationarity is likely to exist, for which a NHPP model is likely to perform better. As the sewer system properties are known to affect blockage rates, it seems likely that these properties also affect system deterioration and hence the Laplace test score. Figure 38, in which all blockage types (i.e. grouping manholes, gully pots and pipes) are considered simultaneously, shows that there is no noticeable influence from diameter, slope or total pipe length on the Laplace test score.

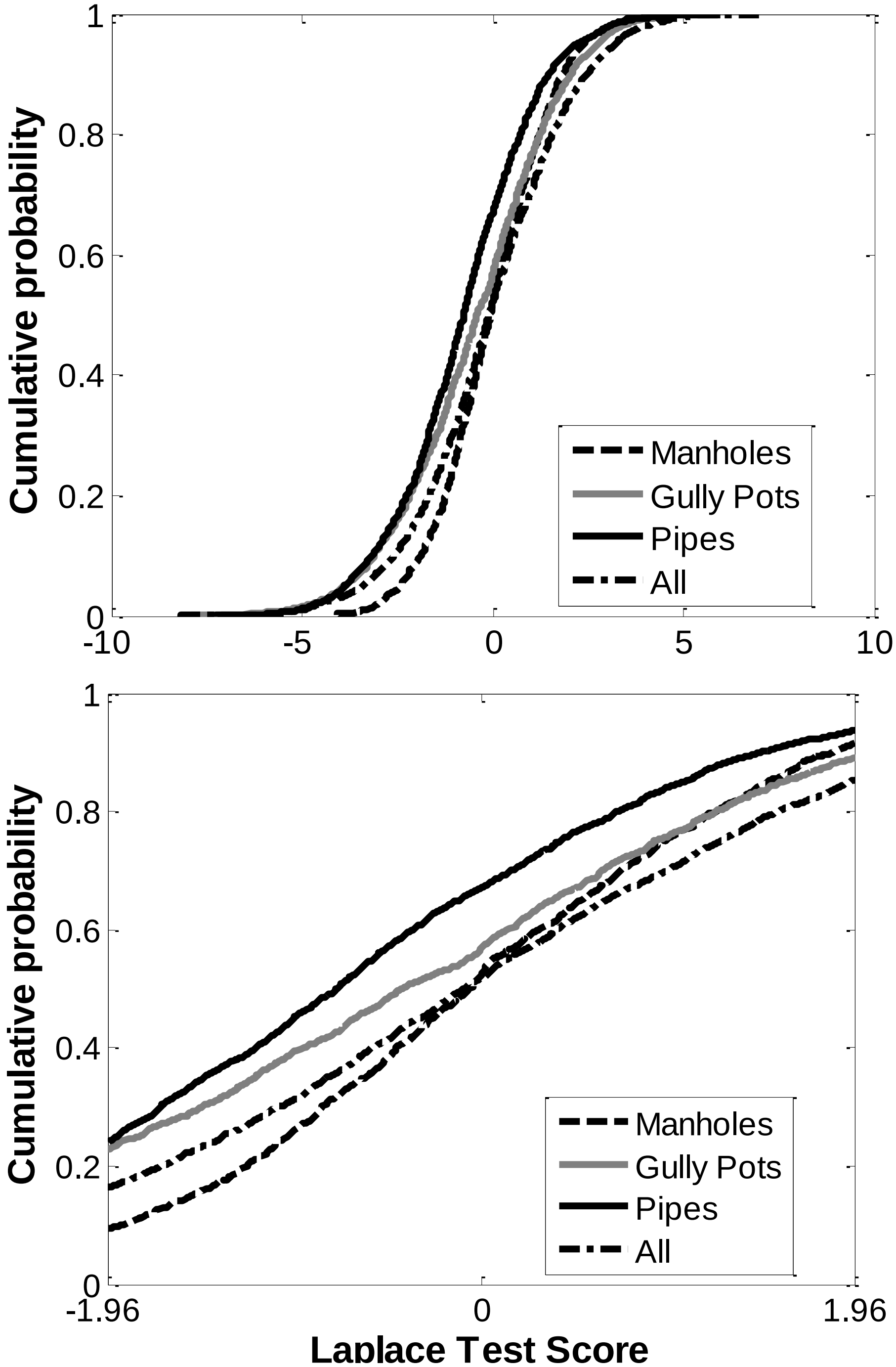


Figure 37. Cumulative probability of the Laplace test score

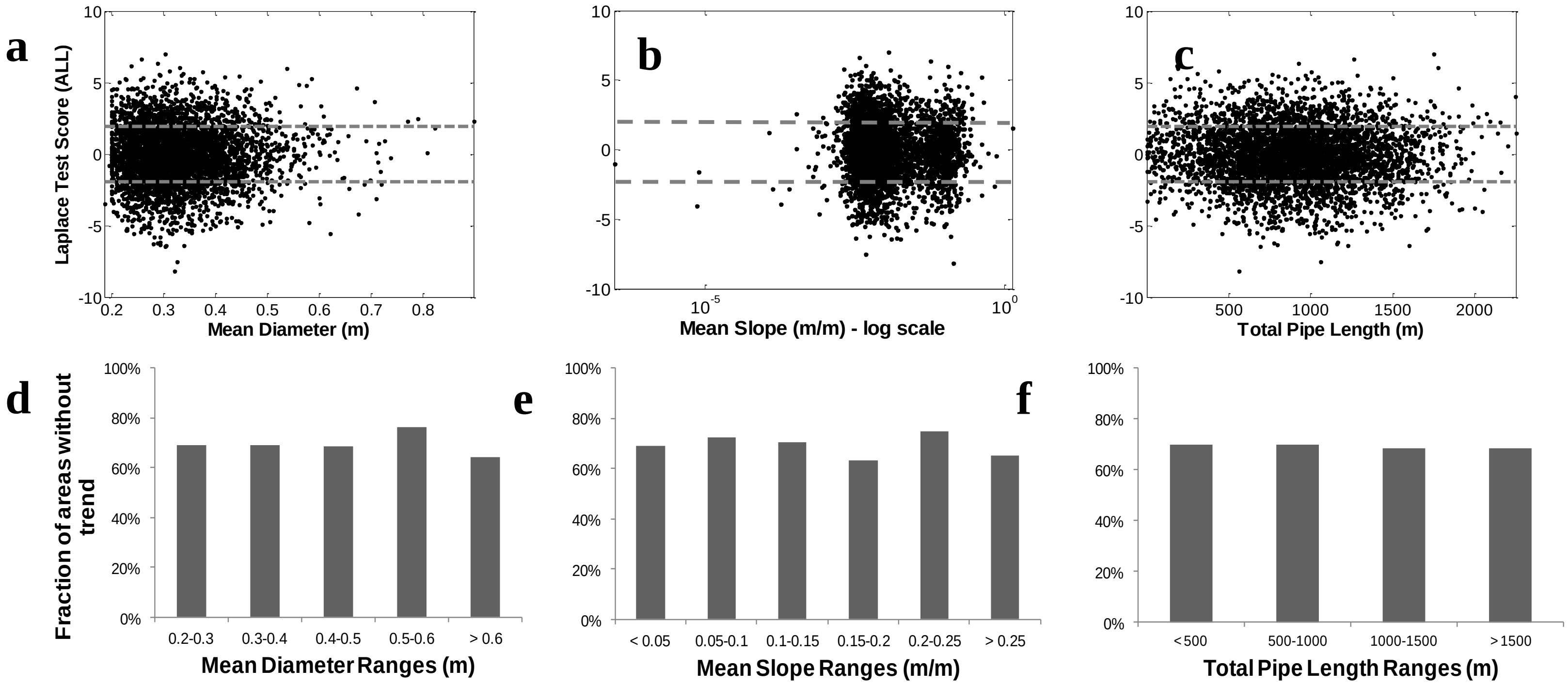


Figure 38. (a) to (c) relationships between sewer system physical properties and estimated Laplace test score. Horizontal lines are confidence intervals within which results are not significant at the 95% level. (d) to (e) fraction of squares without trend for different diameter, slope and total pipe length ranges

6.4.3 Stationary models of blockage dynamics

HPP models

Those squares for which the Laplace test statistic lay between -1.96 and 1.96 were selected for analysis assuming a stationary distribution of inter-arrival times. The agreement with an exponential distribution (HPP) was tested in order to assess whether a more general RP distribution (Weibull or gamma) is appropriate. Following Ascher and Hansen's framework, the total time on test (TTT) statistic was used (Epstein and Sobel 1953) (see Figure 39). When all types of units were considered together, the exponential distribution was implied by the test to be adequate in 35% of the analysed squares. This percentage is 95%, 85% and 73% respectively if manholes, pipes and gully pots are considered separately. Theoretically speaking, if different independent Poisson processes are superimposed, the pooled output is a Poisson process (Cox and Lewis 1966). Results therefore suggest that either blockage in manholes, pipes, gully pots are not independent processes or that the reduced number of samples after disaggregating the blockages into unit types does not support identification of a two-parameter distribution. Further analysis showed that the identifiability of a two-parameter RP model over a HPP model was more related to sample size than any physical difference between the squares.

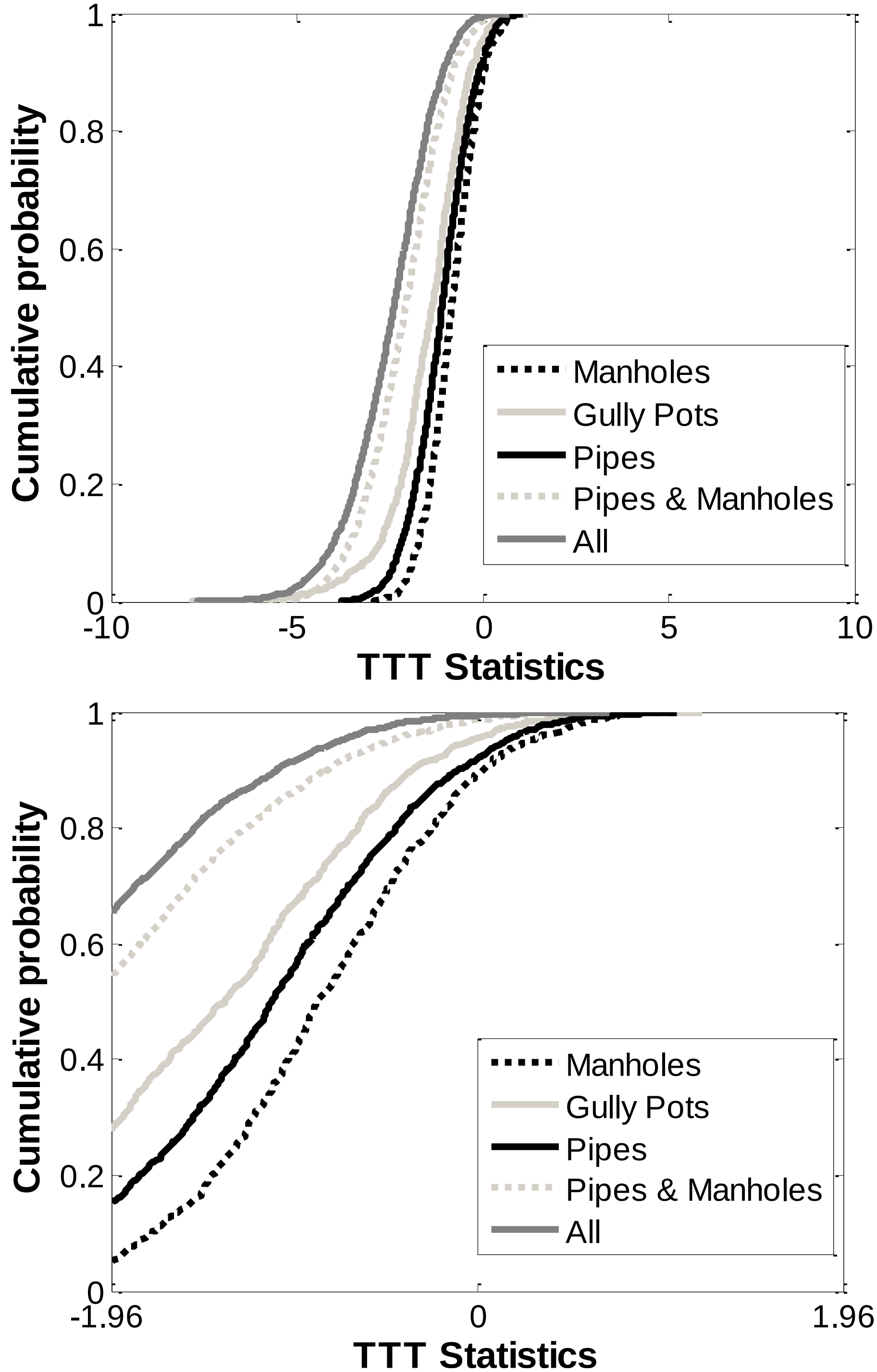


Figure 39. Cumulative probability of the TTT statistic

Aiming to make the results useful in practice, both a stepwise regression and an EPR MOGA based analysis of the HPP model parameter (λ for the exponential distribution when all types of units are bunched together) against system properties was performed. The main idea is to assess if it is possible to estimate model parameters for sites where few or no complaints data are available. Visual inspection of data suggests that there is a slight decrease in λ for increase in pipe length. The stepwise analysis found that λ is, at the 95% confidence level, partly explained by mean pipe diameter, total pipe length, the number of manholes and the COST-S factor. However, the effects are unidentifiable, with R^2 values of 0.05. The EPR MOGA technique could not find any considerably stronger predictive model.

RP models

Where there is enough statistical evidence to reject the exponential distribution (still using only those squares with no apparent trend), there is a more rapid decrease of inter-arrival time after t = 0 than can be described by an exponential (i.e. in all cases the TTT statistic was founded to be < than -1.96 instead of > than 1.96). This points to the use of a Weibull distribution with shape parameter less than 1.0. The Anderson-Darling test (Anderson and Darling 1952) was applied to assess the suitability of using the Weibull distribution to model the observed inter-arrival times. Results at the 95% significance level indicated that in 98% of the squares where the HPP model was not accepted, the manhole inter-arrival blockage time can be represented well by a Weibull distribution. The percentages for gully pots, pipes and all types of units correspond to 71%, 93% and 73% respectively. Two particular squares, with TTT statistic values corresponding to -1.97 and -3.25 respectively (analysing all type of unit simultaneously), are used here to exemplify why a Weibull is considered better than an exponential distribution. Figure 40 shows that the Weibull distributions are able to better represent the observed initial high density of shorter inter-arrival times, illustrating that blockage probability is especially high after a blockage repair. This may seem counter-intuitive: if a single repair has any immediate effect on the overall failure rate in a square then it may be expected to be beneficial. However, the result is intuitive in the sense that it implies that blockages come in clusters, perhaps due to sediment-generating activity or weather conditions affecting the whole square (or a group of squares as can be seen when comparing neighbouring squares), or maybe simply because releasing sediment from one part of the system leads to a knock-on blockage in another structural unit. These considerations may be important when implementing a proactive maintenance scheme.

As for HPP models, a stepwise regression and EPR MOGA analysis of each parameter of the Weibull distribution against system properties were performed. Visual inspection of Weibull's scale and shape data suggests that there is an apparent pipe length dependency. The latter was confirmed by means of the stepwise procedure which identified slope, pipe length and the number of gully pots as statistically significant at the 95% level; however, again the relationships were considered too weak to propose an appropriate predictive model.

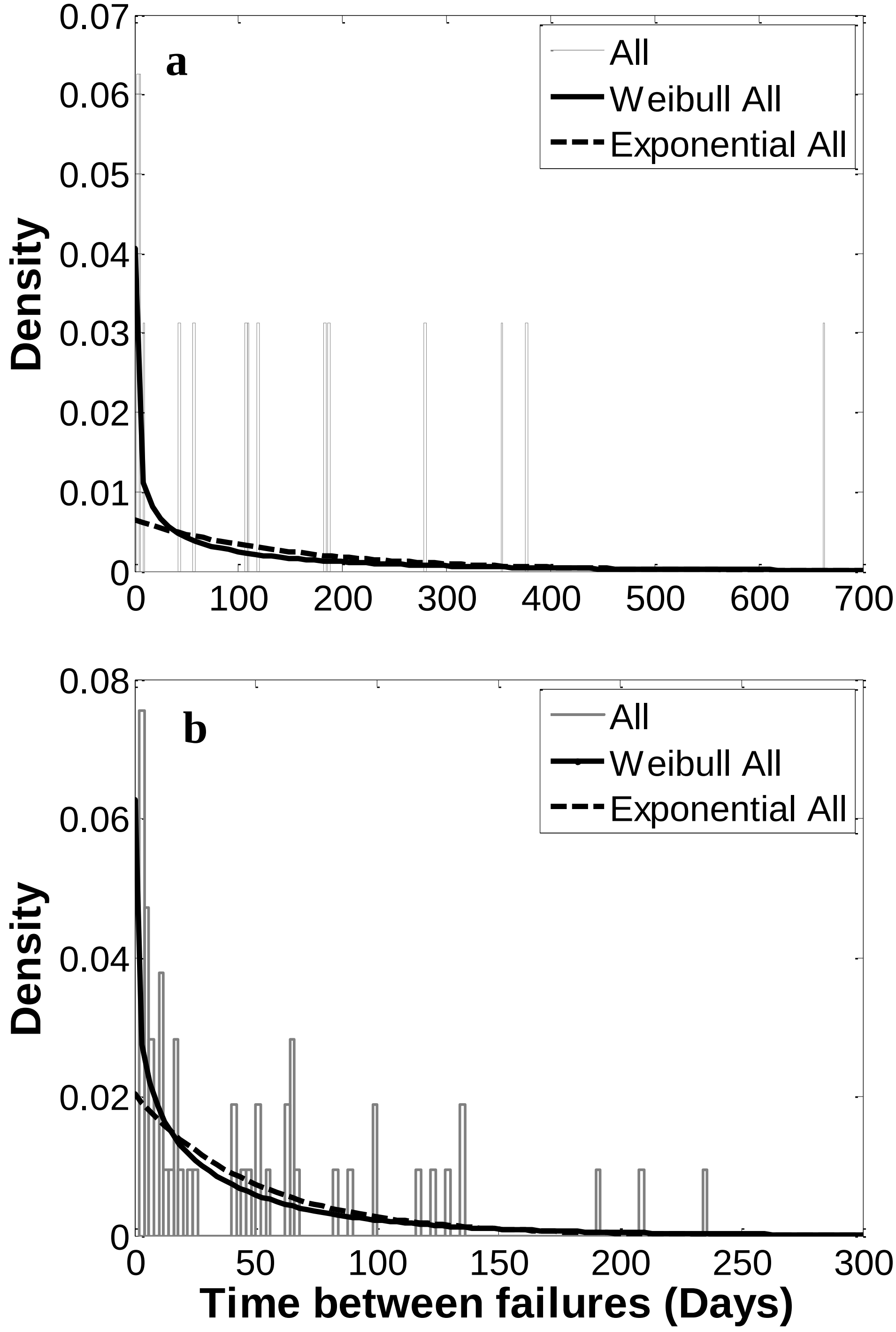


Figure 40. All types of units are simultaneously considered (i.e. pipes, manholes and gully pots). (a) Time between failures (blockages) pdf for a square with TTT statistic value of -1.97 (b) Time between failures (blockages) pdf for a square with TTT statistic value of -3.25 (in both cases bin width correspond to two days)

NHPP-like blockage dynamics

Those squares in which the Laplace trend test found the blockages to be non-stationary are analysed in this sub-section. The Crow model Equation 17 and the Cox and Lewis model Equation 18 were fitted to the blockage inter-arrival times and both gave good agreement about the direction and strength of trend. Results indicate that the reliability of gully pots decreases over time in 67% of those squares, while the corresponding values for pipes and manholes are 79% and 52%. That a significant proportion of the units seem to be improving over time, contrary to the general expectation of gradual deterioration, may be due to random effects in the data or due to fundamental problems being addressed during repairs rather than only cleaning. It is also possible that some squares have temporarily experienced poor waste management practices by building contractors thus increasing the frequency of blockages, and then an improved situation is being observed after the works were finished. When plotting the parameter values in Equation 17 and Equation 18 against physical properties (Figure 41), the only property that seems to have an effect on the strength of trend is the mean slope (this is clearer if the log scale is not used), which was not found to be significant at the 95% level using a stepwise regression procedure (total pipe length and the number of gully pots are shown to have a significant influence).

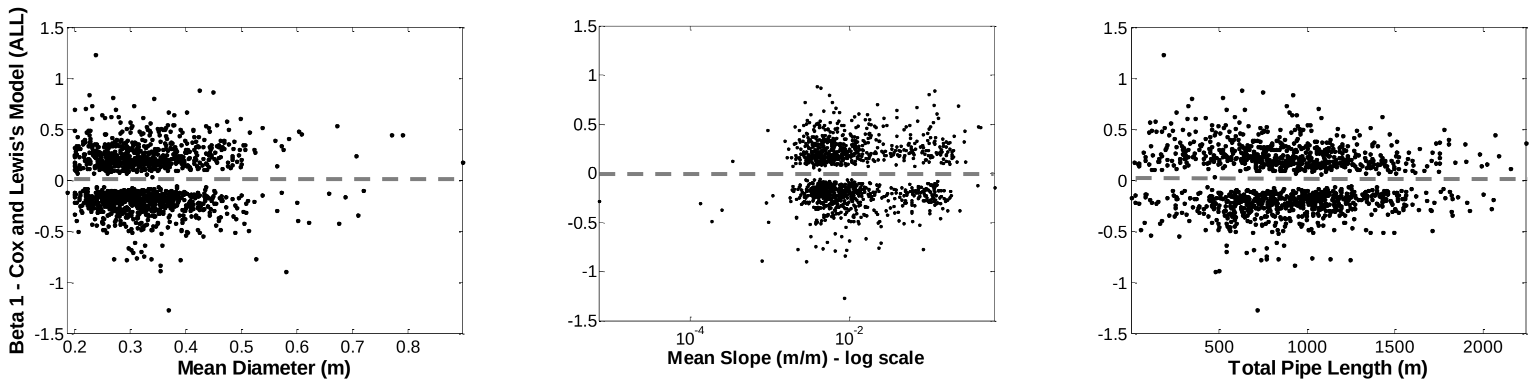


Figure 41. Relationships between sewer system physical properties and estimated Cox and Lewis' model shape parameter

6.4.4 Recommendations for unobserved (or unanalysed) systems

The Bogotá case has shown that complaint databases are one valuable source of information, particularly for sewer system management in large cities, thus reinforcing the need for continuously improving complaint collection platforms where they are in place or to set-up appropriate gathering systems instead. The blockage data analysis framework proposed in this work is considered to be easily transferable to other cases with yet under-exploited failure and/or complaint data bases. This work intended to draw general conclusions about influential factors on blockage likelihood for use where little or no historical blockage data exist but good information about system properties is available. However, the inability to generate either linear or non-linear regression models that strongly relate diameter, slope and total pipe length to average blockage rates illustrates the complexity in and randomness of sewer blockage mechanisms. This is particularly the case in large scale developing urban centres, such as Bogotá, which experience intense building and infrastructure development activities. In contrast, Savic et al. (2006), Savic et al. (2009) and Ugarelli et al. (2009) reported good performance using relatively simple nonlinear models but for presumably relatively “steady-state” cities in the UK and Norway, where the dynamic and unobserved influences on sewer system performance are less strong. Consequently, there is still a need to identify how well the models of Savic et al. and Ugarelli et al. can be transferred to other case studies (particularly in the challenging context of large growing cities).

To try to better characterise the variability in the preferred statistical model type, the spatial distribution of squares better represented by HPP, RP and NHPP models was also assessed here, however no particular spatial pattern or correlation with mean diameters, slopes and total pipe length could be seen. Additional factors may have to be considered such as location and duration of construction sites and other major sources of sediments, and major repairs and upgrades in the sewer system. Nevertheless, given that statistical models have been identified for almost all (about

90% of) squares, the difficulty of generalising models between squares does not preclude application.

Until better understanding of the blockage phenomena is achieved and customer complaints are more widely systematically stored and processed, water companies require methods capable of: (a) operating within levels of current data availability and quality and (b) including expert judgement instead of local statistical evidence. In this context, Fenner et al. (2007) proposed a case-based reasoning approach when good information is available from a small number of pipes (e.g. 20 pipes). They demonstrated that such a methodology can effectively infer maintenance prioritisation avoiding the need for calibration of more statistically or deterministically oriented models. Their approach assigns indices to the pipes relating to their condition (e.g. stable or deteriorating), performance (e.g. number of previous complaints and blockages) and management outcomes (e.g. intervene or non-intervene). Using this information a case library, which contains a variety of pipes in different environments all with known performance histories, can be developed. Classification by similarity can be used to retrieve management information for unknown pipes, enabling proactive maintenance strategies to be developed based on an assessment of similar past cases (Fenner et al. 2007).

6.5 Summary

In most of the analysed areas in the Bogotá case study the inter-arrival time between blockages can be represented by the homogeneous process, but there are a considerable number of areas (up to 34%) for which there is strong evidence of non-stationarity. In most of these cases, the mean blockage rate increases over time, signifying a continual deterioration of the system despite repairs, this being particularly marked for pipe and gully pot related blockages. The physical properties of the system (mean pipe slope, diameter and pipe length) have a clear but weak influence on observed blockage rates. The inability to generate either linear or non-linear regression models that strongly relate diameter, slope and total pipe length to average blockage rates illustrates the complexity in and randomness of sewer blockage mechanisms. This implies dynamic influences for which data do not general exist in large scale applications. The Bogotá case study illustrates the potential value of customer complaints databases and formal analysis frameworks for proactive sewerage maintenance scheduling in large cities.

CHAPTER 7:
CONCLUSIONS AND LOOKING FORWARD

7.1 Chapter 3: The Bogotá case study

Research results have shown that there is a need to consider all the sub-systems of the urban water cycle as one entity when considering pollution control objectives (Lijklema et al. 1993). In this way, the dynamic and complex interactions between the individual sub-systems and the impact of the whole system on the water resources can be investigated. Therefore, effective planning and operation demands a shift from the "fragmented" approach into an "integrated" one which includes the development and implementation of measurement programs and modelling tools at different levels of detail, such as those proposed in this Thesis.

Urban drainage studies, in general, suffer from lack of long term time series due to the high resource requirements of measurement campaigns. This is marked in developing country cities such as Bogotá. Most of the monitoring programmes described in Chapter 3 of this Thesis, and particularly those in which the author of this dissertation was involved, were a compromise between the length of the measurement campaign (constrained by the limited economical and technical resources) and the general aim to appropriately characterise the spatial variability of the sewer system. To achieve such a balance, a wise selection of the monitoring sites was needed to optimise their value for the modelling. This led to monitoring 150 sites, in most of the cases over one 24-hour period using 3-hour composite samples. However, modelling results showed that this level of data acquisition seems to be enough to characterise DWF variability within a large and complex sewer system. Hence the active involvement of the modeller in the field data collection contributed towards the characterisation of the temporal and spatial dynamics of the sewer system as presented in Chapter 4 and 5.

The benefits of having modellers closely engaged with field programmes is being increasingly recognised (Freni and Mannina 2012).

As stated in Chapter 3, this PhD dissertation can be seen as complementary to other recent monitoring and modelling programmes carried out by different research groups at leading Bogotá universities. Such programmes are greatly contributing to a better understanding of the integrated Bogotá system components. However, despite this recent progress, practical application of sustainable solutions is still limited in the context of Bogotá. Further efforts are needed to unify databases, link data analysis and modelling tools and simultaneously monitor the behaviour of the integrated system particularly during highly dynamic conditions (i.e. wet weather conditions) when underlying processes are more difficult to capture and characterise.

7.2 Chapter 4: Characterisation of dry weather loads to sewers

Inclusion of wastewater quality processes in sewer system modelling came together with progress on integrated urban drainage modelling, and this is allowing the development and application of many software tools. However, in general there is a lack of water quality data to support such applications. Chapter 4 proposed and evaluated a model that generates time series of dry weather sub-catchment outflows and water quality for use as inputs to simulations of the larger sewerage system. Recall that the model follows common practice in integrated urban drainage modelling: the DWF patterns are defined by a mean value and a set of hourly multiplier factors. Consequently, the model corresponds to an empirical analysis, rather than an analysis of contributions to the sub-catchment mass balance. Therefore, the proposed methods strongly depend on the quality of available data sets and additional monitoring is recommended to further validate the generator. Input data required for implementing the methodology include (*a*) flow and water quality time series obtained from monitoring campaigns and (*b*) the number of units of each water use component for each sub-catchment.

7.2.1 Particular considerations for sub-model a

This work identified linear and non-linear stochastic models to estimate mean flows and pollutant concentrations, and their uncertainty, based on sub-catchment properties. Even though the models could be further improved if new data are collected, and ideally would be tested using other case studies to evaluate potential for transferability, they are considered useful for Bogotá, and potentially for similar sewer systems where few wastewater quality data are available but information exists about sub-catchment properties.

7.2.2 Particular considerations for sub-model d

In general, 3rd order Fourier series gave satisfactory fits to the diurnal measurements while the 2nd order series did not, and no further improvement was achieved if 4th order series were used. The model was not used to represent DWF patterns over scales longer than one day as such patterns are not visible in the case study data. If weekly or seasonal patterns are identifiable in other cases, these could be incorporated using Fourier series of higher orders. Extending the time domain of model beyond 24 hours would, however, create the additional challenge of estimating additional coefficients in the ungauged site model.

The most successful ungauged site model in terms of fitting observed data required an independent regression model for each of the 24 time-steps Equation 2, disaggregation B). However, constraining the time-profile of each of the five water use components to Fourier series (disaggregation A) was preferred because it provided acceptable performance while being significantly more parsimonious, with reduced uncertainty in its coefficients estimates.

As well as applying the Fourier series to generate the global data and then disaggregating into water user types, Equation 2 was applied directly to the observed data. It was found that such option also explained the hourly variability, but led to confidence bounds that were unnecessarily wide (i.e. the 95 % bounds obtained from the jack-knife validation test included more than 95 % of the observed data for all validation sites). Therefore, although there are other approaches that might be more applicable in other catchments and should be tested, the generation of the global series using a Fourier series Equation 1 followed by also a Fourier series-based disaggregation Equation 2 was considered the best option for generalising to ungauged sites in the Bogotá case.

Regarding the stochastic model, it was assumed that the error mean and variance are uniform over all sites, the serial dependence of the observed errors is described by a

1st order linear autoregressive model and is uniform over sites, and the error dependencies between determinants can be represented by a multi-variate Gaussian model. Some discrepancies between the observations and the assumptions used in the model were identified. In particular, there was serial dependence at lags beyond 1 hour, associated with using 3-hour average values, which are not captured by the model. However, the jack-knife validation procedure suggested that they are not large enough to preclude useful application of the model. To supplement the study of hourly to 3-hour average data, a 3rd order Fourier series model was also fitted to 1-minute 'global' data from another site in Bogotá, and to 5-minute 'global' data from Linz in Austria, showing that the Fourier series and error model worked equally well. These results show that the same mixed deterministic-stochastic model can characterise DWF variables in other gauged sites, although further tests would be required to explore wider applicability to ungauged sites. Further testing, using case studies with multiple gauged sites is recommended.

7.3 Chapter 5: Modelling large scale sewer systems

This work has addressed three primary challenges of integrated assessment of large and complex sewer systems: the need for monitoring data that provides spatial and temporal information for model identification; the need to estimate flow and pollution load inputs including for ungauged sub-catchments; and the identification of a sewer system model with appropriate spatial and temporal resolution, and complexity of process representation.

Implementation of detailed physically based large scale distributed sewer models is often unjustified and unfeasible in developing mega-cities such as Bogotá. Despite recent efforts to develop methods to generate the required observed data Blumensaat et al. (2012), further research is still needed. However, even with relatively good observations, previous research has demonstrated that using detailed sewer system models may not provide advantages over using simplified approaches (Freni et al. 2008). The appropriate level of simplification (in terms of spatial resolution of the model and transformation process representation), however, has not been investigated extensively, especially when the model is required to generate inputs to WWTP models and untreated wastewater discharges at multiple locations over large spatial scales.

Using the City Drain framework (Achleitner et al. 2007), this work demonstrated that it is possible to use semi-distributed conceptual models to represent the sewer system. Furthermore, it was shown that in most cases the apparent random variability in DWF profiles does not allow identification of in-sewer transformation processes thus indicating that in general there is no added value if those complex processes are represented. Instead, uncertainty in DWF profiles is largely controlled by pollutant load variability at the sub-catchment scale. In cases where backwaters caused particularly high retention times, the model did require pollutant transformations and sedimentation/resuspension to be represented. Thus the combination of the monitoring programme, the stochastic load generator and the City Drain modelling framework identified the suitable level of complexity for difference modelling aims.

7.4 Chapter 6: Proactive maintenance of sediment-related sewer blockages

The potential for using a customer complaints databases to develop tools to support prioritisation of maintenance actions in urban drainage systems was demonstrated. Sediment-related blockage data from a customer complaints database in Bogotá were introduced and analysed. The database is exceptionally long (7.5 years), spatially extensive (more than 248,000 failure reports covering a population of 7.5 million) and spatially detailed (data were recorded for each individual 0.03 km^2 square).

In most of the analysed squares (69% of squares when failures of all unit types were modelled together) the inter-arrival time between blockages can be represented by means of stationary renewal process distributions. In general, where enough samples exist to identify a 2-parameter distribution, the Weibull distribution was considered to be best, otherwise a 1-parameter exponential (HPP) distribution was suitable. All the fitted Weibull distributions had a shape parameter less than 1.0, showing that blockage probability tends to be especially high after a blockage and then reduces quickly. This implies that blockages come in clusters due to, for example, construction activity or storm events.

There is still a considerable fraction of squares for which non-stationarity in the inter-arrival time exists, for which a NHPP model performs better. There is no noticeable influence from diameter, slope or total pipe length on the presence of trend. In most of the squares holding trend, the mean blockage rate increases over time, signifying a continual deterioration of the system despite repairs, this being particularly marked for pipe and gully pot related blockages (79% and 69% of squares show deteriorating performance for pipe and gully pots respectively).

Squares with average pipe diameter between 0.2 and 0.4 m, mild slopes (between 10^{-3} and 10^{-2} m/m) and large total pipe length tend to suffer the highest blockage rates not only for pipes but also for manholes and gully pots. However, the inability to generate either linear or non-linear regression models that strongly relate diameter, slope and total pipe length to average blockage rates illustrates the complexity in and randomness of sewer blockage mechanisms. Recall that Bogotá's historical complaints database lacks a link between any particular hydraulic structure and failure properties, instead the link is between square-averaged hydraulic structure properties and failure properties. This averaging inevitably loses some information about failure causes, thus making generalisation more difficult. Furthermore, additional explanatory variables, such as structural conditions, high resolution spatially distributed rainfall data, construction activity and detailed water consumption rates, would ideally be included in an attempt to better describe observed blockage rates.

7.5 A look to the future

In both, developed and developing countries, the major objectives for urban drainage remain public hygiene, flood protection and more recently, environmental protection. In developed countries most of the recent attention has been particularly placed on pollution control to protect the environment while in developing countries, hygiene and flood protection are still the major issues.

A sound approach to urban drainage management should be flexible, interdisciplinary and based on local characteristics. This PhD Thesis has addressed some technical challenges in the context of mathematical model development to support sewer system management in developing country mega-cities. Chapter 1 of this dissertation highlighted the need of an integrated management approach (i.e. simultaneously considering the interactions among the sewer system, treatment facilities and receiving water courses). Science and research can provide a better understanding of integrated system characteristics and its effects on ecosystems and water resources. Regarding sewer system modelling, the author of this Thesis believes that further research efforts, aiming at successful application of integrated models to large urban centres, should be focused on improving our understanding of sewer system dynamics under wet weather conditions.

It is known that the wastewater inflow rate entering at the WWTP depends on the characteristics of the area and population served by the sewer system discharging to the plant and the occurrence of rainfall events. Urban wet weather effluents could induce WWTP malfunction. Under steady-state conditions, a WWTP usually has a satisfactory performance because these conditions are similar to design conditions. However, hydraulic and pollutant loads entering to the waste water treatment plant can be highly variable. This dynamic condition imposes some difficulties in complying with treatment requirements. Therefore, the ability to predict input loads to a treatment facility both for dry and wet weather conditions is very beneficial for the optimization of the treatment process (El-Din and Smith 2002). However, in general, databases to support successful flows and pollutant concentrations modelling during rainfall events particularly for large scale and complex sewer system are not available.

The high complexity of the stormwater quality-related physical phenomena is still not matched by current knowledge or quality of data (Dotto et al. 2011) thus currently precluding large scale and spatially distributed application. Therefore, further research is needed to carry out model-based performance analysis of the integrated urban drainage system under rapid changes in the inflows of mixed wastewater and stormwater and the continuing inflow of both generic (i.e. BOD, COD, TSS) and emerging contaminants (e.g. endocrine disrupters). For this, monitoring programmes are needed to spatially characterise the contribution to pollutant loads at the sub-catchment scale during rainfall events.

In the context of Bogotá, there is available some long-lasting high temporal resolution datasets from two experimental sub-catchments within the sewer system. Using one of these experimental sub-catchments as case study, the author of this dissertation has contributed to the implementation and testing of different modelling approaches to represent the build-up and wash-off of pollutants at the sub-catchment scale (Rodriguez et al. 2010; Manz et al. 2012). Particularly, Rodriguez et al. (2010) implemented and tested computational routines for representing sediment and pollutant loads in order to evaluate catchment surface pollution within the City Drain framework. The tested models estimate the accumulation, erosion and transport of pollutants - aggregately - on urban surfaces and in sewers. Different numerical approaches were tested for their ability to calibrate to the sediment transport conditions. In general, it was observed that using more detailed models for representing pollutant accumulation do not necessarily lead to better results.

Complementary to the work of Rodriguez et al. (2010), Manz et al. (2012) investigated the impact of altering rainfall temporal resolution on the source and magnitude of uncertainties associated with the rainfall model inputs, the hydrological modelling and the pollutant accumulation and wash-off model structure and parameters. The main conclusions of this study can be summarised as follows: (a) asymptotic accumulation models (i.e. those that assume that the surface pollutant load builds up over the antecedent dry days according to a non-linear law) outperformed linear models or those assuming infinite sediment supply; and wash-off models based on rainfall intensity performed better than those based on flow rate and are therefore most recommendable; (b) comparison of the uncertainty band width suggests that 80% of the overall uncertainty in TSS concentration predictions originates from water quality modelling, whereas only 20% stems from hydrological modelling (including rainfall input uncertainties); (c) model prediction uncertainty is sensitive to time resolution and choice of wash-off model; (d) the use of automatic samplers is in general not recommended for wet weather-related studies, as they only offer a limited amount of observations, resulting in uncertainties regarding the increasing limb and the peak of the pollutograph. Furthermore, this leads to skewed results in model performance and uncertainty assessment. The use of high-resolution online devices using surrogate parameters (e.g. turbidity) is more suitable to develop improved models and more robust uncertainty assessment as it may offer a larger abundance of data. This finding reiterates that, while different urban water quality models and calibration methods are available and perform comparably well, when sufficient data is available, the major limiting factor in urban water quality modelling is the scarcity of observations. However, results from these works are not allowing yet an adequate spatial representation of the studied phenomena. Stormwater quality model development is, in general, an area of further research (Dotto et al. 2011). It has to be noted that any further model development should be parsimonious, include appropriate methods for uncertainty quantification and allow characterisation of ungauged sub-catchments based on a limited number of monitoring sites.

NOTATION

$a_{0_{i,j}}$	constant
$a_{z_{i,j}}$	constants
A_i	matrix
$b_{z_{i,j}}$	constants
B_i	matrix
$C_j(t)$	component of $X_{i,j}(t)$ arising from each unit of commercial water use
co_i	known number of commercial units
ε_j	random normal variate
ε_t	sample of modelled values at time t
H	number of inhabitants
I	particular site
$I_j(t)$	component of $X_{i,j}(t)$ arising from each unit of industrial water use
I	percentage of industrial units
i'	percentage of industrial water use consumption
in_i	number of industrial units
J	particular determinant
K	sub-catchment travel time
L	catchment hydraulic length
$M(t)$	component of $X_{i,j}(t)$ arising from each unit of multi-user water use
mu_i	number of multi-user units
M_{0_i}	matrix
M_{1_i}	matrix
M	percentage of multi-user units
NSE^{prob}	Nash-Sutcliffe efficiency adjusted for probabilistic predictions
N_t	number of events

$O_j(t)$	component of $X_{i,j}(t)$ arising from each unit of office-use water use
of_i	number of office-use units
$R_j(t)$	component of $X_{i,j}(t)$ arising from each unit of residential water use
re_i	number of residential units
R	percentage of residential units
T	total number of time-steps
T'	interval from a time origin to the next event
$Variance_{R_j}$	error variance associated with a unit of residential water component
$Variance_{I_j}$	error variance associated with a unit of industrial water component
$Variance_{C_j}$	error variance associated with a unit of commercial water component
$Variance_{O_j}$	error variance associated with a unit of office-use water component
$Variance_{M_j}$	error variance associated with a unit of multi-user water component
$w_{i,j}$	fundamental frequency
${X_{i,j}}^{Obs}(t)$	normalised observed data
$X_{i,j}(t)$	global time series
$X'_{i,j}(t)$	random realisation
x_t^0	observed value
β_0	scale parameter
β_1	growth parameter
Λ	mean rate of occurrence of failure events
ρ_K	serial correlation for lag K
ρ_1	lag 1 serial correlation coefficient

REFERENCES

Achleitner, S., S. De Toffol, C. Engelhard and W. Rauch (2005). "Model based hydropower gate operation for mitigation of CSO impacts by means of river base flow increase." *Water science and technology* 52(5): 87-94.

Achleitner, S., S. Fach, T. Einfalt and W. Rauch (2009). "Nowcasting of rainfall and of combined sewage flow in urban drainage systems." *Water science and technology* 59(6): 1145-1151.

Achleitner, S., M. Möderl and W. Rauch (2007). "CITY DRAIN © - An open source approach for simulation of integrated urban drainage systems." *Environmental modelling & software* 22(8): 1184-1195.

Achleitner, S., M. Möderl and W. Rauch (2007). "Urine separation as part of a real-time control strategy." *Urban Water Journal* 4(4): 233-240.

Achleitner, S. and W. Rauch (2007). "Increase of River Base Flow by Hydropower Gate Operation for Mitigation of CSO Impacts – Potential and Limitations." *Water resources management* 21(9): 1487-1503.

Ahnert, M., V. Kuehn and P. Krebs (2005). "Identification of overall degradation in sewer systems from long-term measurements and consequences for WWTP simulations." *Water science and technology* 52(3): 153-161.

Alex, J., G. Langergraber, M. Ahnert, D. J. Durrenmatt, C. Keysers, M. Reichel and S. Schmuck (2011). Applications for a semi mechanistic dry weather influent model. 8th IWA Symposium on System Analysis and Integrated Assessment – Watermatex 2011. San Sebastian, Spain.

Almeida, M. C., D. Butler and J. W. Davies (1999). "Modelling in-sewer changes in wastewater quality under aerobic conditions." *Water science and technology* 39(9): 63.

Almeida, M. C., D. Butler and E. Friedler (1999). "At-source domestic wastewater quality." *Urban Water* 1(1): 49-55.

Almeida, M. C., D. Butler and J. S. Matos (2000). "In-sewer biodegradation study at the Costa do Estoril interceptor system." *Urban Water* 2(4): 327.

Anderson, T. W. and D. A. Darling (1952). "Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes." *The Annals of Mathematical Statistics* 23(2): 193-212.

Andres-Domenech, I., J. C. Munera, F. Frances and J. B. Marco (2010). "Coupling urban event-based and catchment continuous modelling for combined sewer overflow river impact assessment." *Hydrology and earth system sciences* 14(10): 2057-2072.

Arthur, S., H. Crow and L. Pedezert (2008). "Understanding blockage formation in combined sewer networks." *Proceedings of the Institution of Civil Engineers-Water Management* 161(4): 215-221.

Arthur, S., H. Crow, L. Pedezert and N. Karikas (2009). "The holistic prioritisation of proactive sewer maintenance." *Water science and technology* 59(7): 1385-1396.

Ascher, H. E. and C. K. Hansen (1998). "Spurious exponentiality observed when incorrectly fitting a distribution to nonstationary data." *Ieee Transactions on Reliability* 47(4): 451-459.

Ashley, R. M., A. Fraser, R. Burrows and J. Blanksby (2000). "The management of sediment in combined sewers." *Urban Water* 2(4): 263-275.

Bachoc, A. (1992). "Location and general characteristics of sediment deposits into man-entry combined sewers." *Water science and technology* 25(8): 47-55.

Barnes, J. W. and F. A. Rowe (1978). "Modeling sewer flows using time series analysis." *Journal of the Environmental Engineering Division-Asce* 104(4): 639-646.

Barth, F. and J. Fawell (2001). "The water framework directive and European water policy." *Ecotoxicology and environmental safety* 50(2): 103-105.

Bauwens, W., P. Vanrolleghem and M. Smeets (1996). "An evaluation of the efficiency of the combined sewer- Wastewater treatment system under transient conditions." *Water science and technology* 33(2): 199.

Bauwens, W., P. Vanrolleghem and M. Smeets (1996). "An evaluation of the efficiency of the combined sewer - Wastewater treatment system under transient conditions." *Water science and technology* 33(2): 199-208.

Bechmann, H., M. K. Nielsen, H. Madsen and N. Kjølstad Poulsen (1999). "Grey-box modelling of pollutant loads from a sewer system." *Urban Water* 1(1): 71-78.

Beck, M. B. (1976). "Dynamic modelling and control applications in water quality maintenance." *Water Research* 10(7): 575.

Beck, M. B., A. J. Delayed, P. Lessard, B. A. Finney and L. Simon (1988). Stormwater overflows: modelling impacts on the receiving waters and the treatment plant. Proceedings of a seminar organized by the IAWPRC/IARH Sub-Committee for Urban Runoff Quality Data, as part of the IAWPRC 14th Biennial Conference. Brighton, U.K.

Beenen, A. S., J. G. Langeveld, H. J. Liefting, R. H. Aalderink and H. Velthorst (2011). "An integrated approach for urban water quality assessment." *Water science and technology* 64(7): 1519-1526.

Benedetti, L. (2006). Probabilistic design and upgrade of wastewater treatment plants in the EU Water Framework Directive context, Ghent University, Belgium.

Benedetti, L., D. Bixio and P. A. Vanrolleghem (2006). "Assessment of WWTP design and upgrade options: balancing costs and risks of standards' exceedance." *Water science and technology* 54(6): 371.

Benedetti, L., P. Prat, I. Nopens, M. Poch, C. Turon, B. De Baets and J. Comas (2009). "A new rule generation method to develop a decision support system for integrated management at river basin scale." *Water science and technology* 60(8): 2035-2040.

Berthouex, P. M., W. G. Hunter, L. C. Pallesen and C. Y. Shih (1975). "Modeling sewage treatment plant input BOD data." *Journal of the Environmental Engineering Division-Asce* 101(NEE1): 127-138.

Bertrand-Krajewski, J. and M. Muste (2008). Understanding and managing uncertainty. *Data requirements for integrated urban water management*. T. D. Fletcher and A. Deletic, UNESCO - Taylor & Francis. 1.

Bin Ali, M. T., K. V. Horoshenkov and S. J. Tait (2011). "Rapid detection of sewer defects and blockages using acoustic-based instrumentation." *Water science and technology* 64(8): 1700-1707.

Black&Veatch (2010). 2010 Stormwater utility survey.

Blumensaat, F., M. Wolfram and P. Krebs (2012). "Sewer model development under minimum data requirements." *Environmental Earth Sciences* 65(5): 1427-1437.

Borchardt, D. and F. Sperling (1997). "Urban stormwater discharges: Ecological effects on receiving waters and consequences for technical measures." *Water science and technology* 36(8-9): 173-178.

Breinholt, A., F. O. Thordarson, J. K. Moller, M. Grum, P. S. Mikkelsen and H. Madsen (2011). "Grey-box modelling of flow in sewer systems with state-dependent diffusion." *Environmetrics* 22(8): 946-961.

Bulygina, N., N. McIntyre and H. Wheater (2009). "Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis." *Hydrology and earth system sciences* 13(6): 893-904.

Butler, D. and N. J. D. Graham (1995). "Modeling dry weather wastewater flow in sewer networks." *Journal of Environmental Engineering-Asce* 121(2): 161-173.

Butler, D. and M. Schuetze (2005). "Integrating simulation models with a view to optimal control of urban wastewater systems." *Environmental modelling & software* 20(4): 415.

Calabro, P. S., G. Mannina and G. Viviani (2009). "In sewer processes: mathematical model development and sensitivity analysis." *Water science and technology* 60(1): 107-115.

Camacho, L. A., M. Díaz-Granados and E. Giraldo (2002). Contribución al desarrollo de un modelo de calidad del agua del Río Bogotá (Colombia). XX Latinamerican Congress on Hydraulic (IARH).

Camacho, L. A. and M. A. Díaz-Granados (2003). Metodología para la obtención de un modelo predictivo de transporte de solutos y de calidad del agua en ríos – Caso Río Bogotá. Seminario Internacional La Hidroinformática en la Gestión Integrada de los Recursos Hídricos**:** 73 – 82.

Candela, A., G. Freni, G. Mannina and G. Viviani (2009). "Quantification of diffuse and concentrated pollutant loads at the watershed-scale: an Italian case study." *Water science and technology* 59(11): 2125-2135.

Candela, A., G. Freni, G. Mannina and G. Viviani (2012). "Receiving water body quality assessment: an integrated mathematical approach applied to an Italian case study." *Journal of hydroinformatics* 14(1): 30-47.

Capodaglio, A. G., S. Zheng, V. Novotny and X. Feng (1990). "Stochastic system identification of sewer flow models." *Journal of Environmental Engineering-Asce* 116(2): 284-298.

CAR (2006). Plan de Ordenación y Manejo de la Cuenca Hidrográfica del Río Bogotá, Regional Autonomous Environmental Agency of Cundinamarca.

Caradot, N., D. Granger, J. Chapgier, F. Cherqui and B. Chocat (2011). "Urban flood risk assessment using sewer flooding databases." *Water science and technology* 64(4): 832-840.

Carstensen, J., M. K. Nielsen and H. Strandbaek (1998). "Prediction of hydraulic load for urban storm control of a municipal WWT plant." *Water science and technology* 37(12): 363-370.

Chebbo, G., A. Bachoc, D. Laplace and B. Le Guennec (1995). "The transfer of solids in combined sewer networks." *Water science and technology* 31(7): 95.

Cohen, B. (2006). "Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability." *Technology in Society* 28(1-2): 63-80.

Cox, D. R. and P. A. W. Lewis (1966). *The statistical analysis of series of events*, Methuen & Co Ltd.

Crow, L. H. (1975). Reliability analysis for complex, repairable systems, Army Materiel Systems Analysis Activity – AMSAA. Technical report No. 138.

DANE (2005). Colombian census, National Statistical Technical Department.

De Keyser, W., V. Gevaert, F. Verdonck, B. De Baets and L. Benedetti (2010). "An emission time series generator for pollutant release modelling in urban areas." *Environmental modelling & software* 25(4): 554-561.

De Toffol, S. (2006). Sewer system performance assessment - an indicators based methodology, University of Innsbruck.

De Toffol, S., C. Engelhard and W. Rauch (2007). "Combined sewer system versus separate system- a comparison of ecological and economical performance indicators." *Water science and technology* 55(4): 255.

Deletic, A., C. B. S. Dotto, D. T. McCarthy, M. Kleidorfer, G. Freni, G. Mannina, M. Uhl, M. Henrichs, T. D. Fletcher, W. Rauch, J. L. Bertrand-Krajewski and S. Tait (2012). "Assessing uncertainties in urban drainage models." *Physics and Chemistry of the Earth, Parts A/B/C* 42–44(0): 3-10.

Deletic, A. and T. D. Fletcher (2008). Chapter 2: Overview of guiding principles. *Data Requirements for Integrated Urban Water Management*. A. Deletic and T. D. Fletcher, UNESCO - Taylor & Francis.

Devesa, F., J. Comas, C. Turon, A. Freixó, F. Carrasco and M. Poch (2009). "Scenario analysis for the role of sanitation infrastructures in integrated urban wastewater management." *Environmental Modelling & Software* 24(3): 371-380.

Díaz-Granados, M. A., J. P. Rodríguez, M. S. Rodríguez, J. C. Penagos, L. A. Camacho, S. Achleitner, Č. Maksimović and N. McIntyre (2009). "Towards a paradigm shift in urban drainage management and modelling in developing countries." *Revista de Ingeniería – Universidad de los Andes* 30: 133-150.

DNP (2004). Estrategia para el manejo ambiental del Río Bogotá. Conpes, National Planning Department. Conpes 3320.

Dotto, C. B. S., M. Kleidorfer, A. Deletic, W. Rauch, D. T. McCarthy and T. D. Fletcher (2011). "Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data." *Environmental Modelling & Software* 26(10): 1225-1239.

Duarte, M. P., L. A. Camacho and E. A. Rodríguez (2009). Medición y análisis del fenómeno de primer lavado en cuencas urbanas. Caso: Campus Universidad Nacional de Colombia - Sede Bogotá. Seminario Internacional "Un Nuevo paradigma en el manejo integrado del agua en áreas urbanas" – AGUA 2009. Cali, Colombia.

EAAB (2006). Plan Maestro de Acueducto y Alcantarillado (Documento Técnico Soporte), Bogotá Water Supply and Sewerage Company.

Efron, B. (1979). "Bootstrap methods: another look at the jackknife." *The Anals of Statistics* 7(1): 1-26.

El-Din, A. G. and D. W. Smith (2002). "A neural network model to predict the wastewater inflow incorporating rainfall events." *Water Research* 36(5): 1115.

Engelhard, C. (2006). Assessing the impact of urban drainage measures with regard to the Water Framework Directive, University of Innsbruck.

Engelhard, C., S. De Toffol and W. Rauch (2008). "Suitability of CSO performance indicators for compliance with ambient water quality targets." *Urban Water Journal* 5(1): 43 - 49.

Epstein, B. and M. Sobel (1953). "Life Testing." *Journal of the American Statistical Association* 48(263): 486-502.

Estupiñan, H. A., E. A. Rodríguez and L. A. Camacho (2009). Evaluación de la contaminación por escorrentía urbana y su modelación mediante la simulación de los procesos de acumulación y lavado. Campus Universidad Nacional de Colombia - Sede Bogotá. Seminario Internacional "Un Nuevo paradigma en el manejo integrado del agua en áreas urbanas" – AGUA 2009. Cali, Colombia.

Feeney, C. S., S. Thayer, M. Bonomo and K. Martel (2009). White paper on condition assessment of wastewater collection systems, United States Environmental Protection agency - EPA.

Fenner, R. A. (2000). "Approaches to sewer maintenance: a review." *Urban Water* 2(4): 343-356.

Fenner, R. A., G. McFarland and O. Thorne (2007). "Case-based reasoning approach for managing sewerage assets." *Proceedings of the Institution of Civil Engineers-Water Management* 160(1): 15-24.

Fenner, R. A. and L. Sweeting (1999). "A decision support model for the rehabilitation of "non-critical" sewers." *Water science and technology* 39(9): 193-200.

Fenner, R. A., L. Sweeting and M. J. Marriott (2000). "A new approach for directing proactive sewer maintenance." *Proceedings of the Institution of Civil Engineers-Water Maritime and Energy* 142(2): 67-77.

Flamink, C., J. Langeveld and F. Clemens (2005). "Aerobic transformations in sewer systems: are they relevant?" *Water science and technology* 52(3): 163-170.

Fletcher, T. D., V. Mitchell, A. Deletic and C. Maksimovic (2008). Introduction. *Data requirements for integrated urban water management*. T. D. Fletcher and A. Deletic, UNESCO and Taylor & Francis.

Freni, G., M. Maglionico, G. Mannina and G. Viviani (2008). "Comparison between a detailed and a simplified integrated model for the assessment of urban drainage environmental impact on an ephemeral river." *Urban Water Journal* 5(2): 87 - 96.

Freni, G. and G. Mannina (2010). "Uncertainty in water quality modelling The applicability of Variance Decomposition Approach." *Journal of Hydrology* 394(3-4): 324-333.

Freni, G. and G. Mannina (2012). "The identifiability analysis for setting up measuring campaigns in integrated water quality modelling." *Physics and Chemistry of the Earth, Parts A/B/C* 42–44(0): 52-60.

Freni, G., G. Mannina and G. Viviani (2009). "Assessment of data availability influence on integrated urban drainage modelling uncertainty." *Environmental modelling & software* 24(10): 1171-1181.

Freni, G., G. Mannina and G. Viviani (2009). "Uncertainty assessment of an integrated urban drainage model." *Journal of Hydrology* 373(3-4): 392-404.

Freni, G., G. Mannina and G. Viviani (2011). "Assessment of the integrated urban water quality model complexity through identifiability analysis." *Water Research* 45(1): 37-50.

Friedler, E. and D. Butler (1996). "Quantifying the inherent uncertainty in the quantity and quality of domestic wastewater." *Water science and technology* 33(2): 65-78.

Gamerith, V., G. Gruber and D. Muschalla (2011). "Single- and Multievent Optimization in Combined Sewer Flow and Water Quality Model Calibration." *Journal of Environmental Engineering-Asce* 137(7): 551-558.

Gasperi, J., M. Kafi-Benyahia, C. Lorgeoux, R. Moilleron, M. C. Gromaire and G. Chebbo (2008). "Wastewater quality and pollutant loads in combined sewers during dry weather periods." *Urban Water Journal* 5(4): 305-314.

Gaudoin, O. (1992). "Optimal properties of the Laplace trend test for software reliability models." *Ieee Transactions on Reliability* 41(4): 525-532.

Gerard, C. and B. Chocat (1999). "Prediction of sediment build-up from analysis of physical network data." *Water science and technology* 39(9): 185-192.

Gernaey, K. V., X. Flores-Alsina, C. Rosen, L. Benedetti and U. Jeppsson (2011). "Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach." *Environmental modelling & software* 26(11): 1255-1267.

Giustolisi, O. and D. A. Savic (2006). "A symbolic data-driven technique based on evolutionary polynomial regression." *Journal of hydroinformatics* 8(3): 207-222.

Giustolisi, O. and D. A. Savic (2009). "Advances in data-driven analyses and modelling using EPR-MOGA." *Journal of hydroinformatics* 11(3-4): 225-236.

González, J. D. (2011). Modelación Integrada del Sistema de Drenaje – PTAR – Río de la ciudad de Bogotá. Escenarios de Control Regional, Universidad Nacional de Colombia.

Gruber, G., S. Winkler and A. Pressl (2004). "Quantification of pollution loads from CSOs into surface water bodies by means of online techniques." *Water science and technology* 50(11): 73-80.

Gruber, G., S. Winkler and A. Pressl (2005). "Continuous monitoring in sewer networks an approach for quantification of pollution loads from CSOs into surface water bodies." *Water science and technology* 52(12): 215-223.

Grucon-IEH-Soprin (1999). Estudio para la rehabilitación del sistema de alcantarillado, Bogotá Water Supply and Sewage Company.

Gudjonsson, G., J. Vollertsen and T. Hvitved-Jacobsen (2002). "Dissolved oxygen in gravity sewers - measurement and simulation." *Water science and technology* 45(3): 35-44.

Harremoes, P. (1988). "Stochastic models for estimation of extreme pollution from urban runoff." *Water Research* 22(8): 1017-1026.

Harremoës, P. (2002). "Integrated urban drainage, status and perspectives." *Water science and technology* 45(3): 1.

Henze, N. and B. Zirkler (1990). "A Class of Invariant Consistent Tests for Multivariate Normality." *Communications in Statistics –Theory and Methods* 19(10): 3595–3618.

Hochedlinger, M., P. Hofbauer, G. Wandl, S. Meyer, W. Rauch, H. Kroiss and M. Heindl (2006). Online UV-VIS measurements - The basis for future pollution based sewer real time control in Linz. 2nd International IWA Conference on Sewer Operation and Maintenance. Vienna, Austria.

Hochedlinger, M., H. Kainz and W. Rauch (2005). Assessment of CSO loads - based on UV/VIS spectroscopy by means of different regression methods. 10th International Conference on Urban Drainage. Copenhagen, Denmark.

Huisman, J. L., T. Gasser, C. Gienal, M. Kuhni, P. Krebs and W. Gujer (2004). "Quantification of oxygen fluxes in a long gravity sewer." *Water Research* 38(5): 1237-1247.

Huisman, J. L., P. Krebs and W. Gujer (2003). "Integral and unified model for the sewer and wastewater treatment plant focusing on transformations." *Water science and technology* 47(12): 65.

Hvitved-Jacobsen, T. (1998). "Wastewater quality changes during transport in sewers-an integrated aerobic and anaerobic model concept for carbon and sulfur microbial transformations." *Water science and technology* 38(10): 257.

Hvitved-Jacobsen, T., J. Vollertsen and J. S. Matos (2002). "The sewer as a bioreactor-a dry weather approach." *Water science and technology* 45(3): 11.

Hvitved-Jacobsen, T., J. Vollertsen and P. H. Nielsen (1998). "A process and model concept for microbial wastewater transformations in gravity sewers." *Water science and technology* 37(1): 233.

ifak (2009). SIMBA 6. Reference and user manual, Institut for Automation und Kommunikation e. V. Magdeburg.

Jacobson, C. R. (2011). "Identification and quantification of the hydrological impacts of imperviousness in urban catchments: A review." *Journal of Environmental Management* 92(6): 1438-1448.

Jamwal, P., A. K. Mittal and J. M. Mouchel (2008). "Effects of urbanisation on the quality of the urban runoff for Delhi watershed." *Urban Water Journal* 5(3): 247 - 257.

Jin, Y. L. and A. Mukherjee (2010). "Modeling Blockage Failures in Sewer Systems to Support Maintenance Decision Making." *Journal of Performance of Constructed Facilities* 24(6): 622-633.

Kaijun, W., G. Zeeman and G. Lettinga (1995). "Alteration in sewage characteristics upon aging." *Water science and technology* 31(7): 191.

Kannapiran, A., A. Chanan, S. G., P. Tambosis, J. Jeyakumaran and J. Kandasamy (2007). "An innovative model for sustainable cost effective management of stormwater drainage assets." *Water Asset Management International* 3(2): 4-11.

Karpf, C. and P. Krebs (2011). "Quantification of groundwater infiltration and surface water inflows in urban sewer networks based on a multiple model approach." *Water Research* 45(10): 3129-3136.

Kelleher, C. and T. Wagener (2011). "Ten guidelines for effective data visualization in scientific publications." *Environmental modelling & software* 26(6): 822-827.

Khalil, B., T. Ouarda and A. St-Hilaire (2011). "Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis." *Journal of Hydrology* 405(3-4): 277-287.

Kleidorfer, M., M. Moderl, S. Fach and W. Rauch (2009). "Optimization of measurement campaigns for calibration of a conceptual sewer model." *Water science and technology* 59(8): 1523-1530.

Korving, H., F. Clemens and J. M. van Noortwijk (2006). "Statistical modeling of the serviceability of sewage pumps." *Journal of Hydraulic Engineering-Asce* 132(10): 1076-1085.

Lamprea, K. and V. Ruban (2011). "Pollutant concentrations and fluxes in both stormwater and wastewater at the outlet of two urban watersheds in Nantes (France)." *Urban Water Journal* 8(4): 219-231.

Lan-Anh, N. T., J. Boxall, A. Saul and P. Willems (2012). "Integrated river flow modelling: A case study." *Urban Water Journal*: 1-18.

Langergraber, G., J. Alex, N. Weissenbacher, D. Woerner, M. Ahnert, T. Frehmann, N. Halft, I. Hobus, M. Plattes, V. Spering and S. Winkler (2008). "Generation of diurnal variation for influent data for dynamic simulation." *Water science and technology* 57(9): 1483-1486.

Langeveld, J. G. (2004). Interactions within wastewater systems, Delft University of Technology, the Netherlands.

Langeveld, J. G., F. H. L. R. Clemens and J. H. J. M. van der Graaf (2003). "Interactions within the wastewater system: requirements for sewer processes modelling." *Water science and technology* 47(4): 101.

Laplace, D., A. Bachoc, Y. Sanchez and D. Dartus (1992). "Trunk sewer clogging development - Description and solutions." *Water science and technology* 25(8): 91-100.

Larrarte, F. (2008). "Suspended solids within sewers: an experimental study." *Environmental fluid mechanics* 8(3): 249-261.

Lau, J., D. Butler and M. Schütze (2002). "Is combined sewer overflow spill frequency/volume a good indicator of receiving water quality impact?" *Urban Water* 4(2): 181-189.

Lee, T. R. (2000). "Urban water management for better urban life in Latin America." *Urban Water* 2(1): 71-78.

Lijklema, L. (1995). "Water quality standards: Sense and nonsense." *Water science and technology* 31(8): 321.

Lijklema, L., J. M. Tyson and A. Lesouef (1993). "Interactions between sewers, treatment plants and receiving waters in urban areas - a summary of the Interurba 92 workshop conclusions." *Water science and technology* 27(12): 1-29.

Lilliefors, H. W. (1967). "On the Kolmogorov-Smirnov test for normality with mean and variance unknown." *Journal of the American Statistical Association* 62(318): 399–402.

Lin, Z. L. and M. Beck (2012). "Accounting for structural error and uncertainty in a model: An approach based on model parameters as stochastic processes." *Environmental modelling & software* 27-28: 97-111.

Maksimović, Č. and A. Tejada-Guibert (2001). *Frontiers in urban water management*, IWA Publishing.

Mannina, G. and G. Viviani (2009). "Separate and combined sewer systems: a long-term modelling approach." *Water science and technology* 60(3): 555-565.
Manz, J. B., J. P. Rodríguez, Č. Maksimović and N. McIntyre (2012). Impact of rainfall temporal resolution on urban water quality modelling performance and uncertainties. 9th International Conference on Urban Drainage Modelling. Belgrade, Serbia.
Marlow, D. R., F. Boulaire, D. J. Beale, C. Grundy and M. Moglia (2011). "Sewer Performance Reporting: Factors That Influence Blockages." *Journal of Infrastructure Systems* 17(1): 42-51.
Maryns, F. and W. Bauwens (1997). "The application of the activated sludge model no.1 to a river environment." *Water science and technology* 36(5): 201.
Mashford, J., D. Marlow, D. Tran and R. May (2011). "Prediction of Sewer Condition Grade Using Support Vector Machines." *Journal of Computing in Civil Engineering* 25(4): 283-290.
Massey, F. J. (1951). "The Kolmogorov-Smirnov Test for Goodness of Fit." *Journal of the American Statistical Association* 46(253): 68-78.
Matalas, N. C. (1967). "Mathematical assessment of synthetic hydrology." *Water Resour. Res.* 3(4): 937-945.
MAVDT (2004). Plan Nacional de Manejo de Aguas Residuales Municipales en Colombia, Ministry of Environment, Housing and Territorial Development.
McIntyre, N. (2004). Analysis of uncertaitny in river water quality modelling. Department of Civil and Environmental Engineering, Imperial College London.
McIntyre, N. R. and H. S. Wheater (2004). "Calibration of an in-river phosphorus model: prior evaluation of data needs and model uncertainty." *Journal of Hydrology* 290(1-2): 100-116.
McMichael, F. and F. C. Vigani (1972). "Discussion of “Characterization of time-varying organic loads” by A. T. Wallace and D. M. Zollman." *Journal of the Sanitary Engineering Division – Proceedings of the American Society of Civil Engineers* 98(2): 444-455.
Meirlaen, J. (2002). Immission based real-time control of the integrated urban wastewater system, Ghent University, Belgium.
Mestra, G. L. (2008). Identificación de conexiones erradas en alcantarillados separados mediante la aplicación de herramientas computacionales, Universidad de los Andes, Colombia. M.Sc.
Metadier, M. and J. L. Bertrand-Krajewski (2011). "Assessing dry weather flow contribution in TSS and COD storm events loads in combined sewer systems." *Water science and technology* 63(12): 2983-2991.
Mirkin, B. G. (2005). *Clustering for data mining: a data recovery approach*, Chapman & Hall/CRC.
Mitchell, V. (2006). "Applying Integrated Urban Water Management Concepts: A Review of Australian Experience." *Environmental management* 37(5): 589-605.
Mitchell, V. G., H. Duncan, M. Inman, M. Rahilly, J. Stewart, A. Vieritz, P. Holt, A. Grant, T. D. Fletcher, J. Coleman, S. Maheepala, A. Sharma, A. Deltic and P. Bree (2007). State of the art in integrated urban water modelling. 6th international conference on sustainable techniques and strategies in urban water management Lyon, France.

Mora, O., D. Alvarado and A. Torres (2011). Economical analysis of the application of porous pavements on Bogota roads. 12th International Conference on Urban Drainage – ICUD 2011. Porto Alegre, Brazil.

Muschalla, D., M. Schutze, K. Schroeder, M. Bach, F. Blumensaat, G. Gruber, K. Klepiszewski, M. Pabst, A. Pressl, N. Schindler, A. M. Solvi and J. Wiese (2009). "The HSG procedure for modelling integrated urban wastewater systems." *Water science and technology* 60(8): 2065-2075.

Nielsen, P. H., K. Raunkjaer, N. H. Norsker, N. A. Jensen and T. Hvitved-Jacobsen (1992). "Transformation of wastewater in sewer systems- a review." *Water science and technology* 25(6): 17.

Ort, C., C. Schaffner, W. Giger and W. Gujer (2005). "Modeling stochastic load variations in sewer systems." *Water science and technology* 52(5): 113-122.

Pantsar-Kallio, M., S. P. Mujunen, G. Hatzimihalis, P. Koutoufides, P. Minkkinen, P. J. Wilkie and M. A. Connor (1999). "Multivariate data analysis of key pollutants in sewage samples: a case study." *Analytica Chimica Acta* 393(1-3): 181-191.

Parkinson, J. and O. Mark (2005). *Urban stormwater management in developing countries*, IWA Publishing.

Pfister, A., A. Stein, S. Schlegel and B. Teichgraber (1998). "An integrated approach for improving the wastewater discharge and treatment systems." *Water science and technology* 37(1): 341.

Prat, P., L. Benedetti, L. Corominas, J. Comas and M. Poch (2012). "Model-based knowledge acquisition in environmental decision support system for wastewater integrated management." *Water science and technology* 65(6): 1123-1129.

Price, R. K. and G. J. Catterson (1997). "Monitoring and modelling in urban drainage." *Water science and technology* 36(8-9): 283-287.

Raciny, I. C., J. P. Rodríguez, M. A. Díaz-Granados and J. D. Pérez (2008). Modelación de la Calidad del Agua de los Ríos Urbanos Salitre, Fucha y Tunjuelo. XXIII Latinamerican Congress on Hydraulic (IARH). Cartagena, Colombia.

Rada, A. and A. Torres (2011). Preliminary assessment of the influence of Salitre basin rainfall on the pumping operation of Salitre wastewater treatment plant (Bogota). 12th International Conference on Urban Drainage – ICUD 2011. Porto Alegre, Brazil.

Rauch, W., H. Aalderink, P. Krebs, W. Schilling and P. Vanrolleghem (1998). "Requirements for integrated wastewater models- driven by receiving water objectives." *Water science and technology* 38(11): 97.

Rauch, W. and P. Harremoes (1998). "Correlation of combined sewer overflow reduction due to real-time control and resulting effect on the oxygen concentration in the river." *Water science and technology* 37(12): 69.

Rauch, W., K. Seggelke, R. Brown and P. Krebs (2005). "Integrated approaches in urban storm drainage: Where do we stand?" *Environmental management* 35(4): 396-409.

Rauch, W., P. A. Vanrolleghem, J. L. Bertrand-Krajewski, W. Schilling, O. Mark, P. Krebs and M. Schutze (2002). "Deterministic modelling of integrated urban drainage systems." *Water science and technology* 45(3): 81-94.

Richardson, C. W. (1981). "Stochastic simulation of daily precipitation, temperature, and solar radiation." *Water Resour. Res.* 17(1): 182-190.

Rieckermann, J., J. Anta, A. Scheidegger and C. Ort (2011). "Assessing Wastewater Micropollutant Loads with Approximate Bayesian Computations." *Environmental Science & Technology* 45(10): 4399-4406.

Rodriguez, F., H. Andrieu and F. Morena (2008). "A distributed hydrological model for urbanized areas - Model development and application to case studies." *Journal of Hydrology* 351(3-4): 268-287.

Rodriguez, J. P., S. Achleitner, M. Moderl, W. Rauch, C. Maksimovic, N. McIntyre, M. A. Diaz-Granados and M. S. Rodriguez (2010). "Sediment and pollutant load modelling using an integrated urban drainage modelling toolbox: an application of City Drain." *Water science and technology* 61(9): 2273-2282.

Romanova, A., M. T. Bin Ali and K. V. Horoshenkov (2011). "A non-intrusive acoustic method of sewer pipe survey. A novel, inexpensive and rapid primary survey technique." *AWA Water Journal - Asset Management* 38(4): 103 – 105.

Rossi, L., V. Krejci, W. Rauch, S. Kreikenbaum, R. Fankhauser and W. Gujer (2005). "Stochastic modeling of total suspended solids (TSS) in urban areas during rain events." *Water Research* 39(17): 4188-4196.

Savic, D., S. Djordjevic, G. Dorini, W. Shepherd, A. Cashman and A. J. Saul (2005). "COST-S: a new methodology and tools for sewerage asset management based on whole life costs." *Water Asset Management International* 1(4).

Savic, D., O. Giustolisi, L. Berardi, W. Shepherd, S. Djordjevic and A. Saul (2006). "Modelling sewer failure by evolutionary computing." *Proceedings of the Institution of Civil Engineers-Water Management* 159(2): 111-118.

Savic, D. A., O. Giustolisi and D. Laucelli (2009). "Asset deterioration analysis using multi-utility data and multi-objective data mining." *Journal of hydroinformatics* 11(3-4): 211-224.

Schellart, A. N. A., S. J. Tait and R. M. Ashley (2010). "Towards quantification of uncertainty in predicting water quality failures in integrated catchment model studies." *Water Research* 44(13): 3893-3904.

Schilling, W., W. Bauwens, D. Borchardt, P. Krebs, W. Rauch and P. Vanrolleghem (1997). On the relation between urban wastewater management needs and receiving water objectives. XXVII IAHR Congress. San Francisco.

SDA (2007). Resolución No. 3257, Environmental Secretariat for the District of Bogotá.

Singh, K. P., A. Malik, D. Mohan, S. Sinha and V. K. Singh (2005). "Chemometric data analysis of pollutants in wastewater - a case study." *Analytica Chimica Acta* 532(1): 15-25.

Solvi, A.-M. (2006). Modelling the sewer-treatment-urban river system in view of the EU Water Framework Directive, Ghent University - Belgium. Ph.D.

Szollosi-Nagy, A. (2008). Foreword. *Urban Water Cycle Processes and Interactions* J. Marsalek, B. Jimenez-Cisneros, M. Karamouz, P.-A. Malmquist, J. Goldenfum and B. Chocat, UNESCO and Taylor & Francis.

Tanaka, N. and J. Hvitved-Jacobsen (1998). "Transformations of wastewater organic matter in sewers under changing aerobic/anaerobic conditions." *Water science and technology* 37(1): 105.

ten Veldhuis, J. A. E. and F. Clemens (2010). "Flood risk modelling based on tangible and intangible urban flood damage quantification." *Water science and technology* 62(1): 189-195.

ten Veldhuis, J. A. E. and F. Clemens (2011). "The efficiency of asset management strategies to reduce urban flood risk." *Water science and technology* 64(6): 1317-1324.

ten Veldhuis, J. A. E., F. Clemens, G. Sterk and B. R. Berends (2010). "Microbial risks associated with exposure to pathogens in contaminated urban flood water." *Water Research* 44(9): 2910-2918.

ten Veldhuis, J. A. E., F. Clemens and P. van Gelder (2009). "Fault tree analysis for urban flooding." *Water science and technology* 59(8): 1621-1629.

ten Veldhuis, J. A. E., R. C. Harder and M. Loog (2010). "Automatic classification of municipal call data to support quantitative risk analysis of urban drainage systems." *Structure and Infrastructure Engineering*: 1-10.

Thordarson, F., A. Breinholt, J. Møller, P. Mikkelsen, M. Grum and H. Madsen "Evaluation of probabilistic flow predictions in sewer systems using grey box models and a skill score criterion." *Stochastic Environmental Research and Risk Assessment*: 1-12.

Todeschini, S., S. Papiri and R. Sconfietti (2011). "Impact assessment of urban wet-weather sewer discharges on the Vernavola river (Northern Italy)." *Civil Engineering and Environmental Systems* 28(3): 209-229.

Torres, A., J. L. Estupiñán and H. O. Zapata (2011). Proposal and assessment of rainwater harvesting scenarios on the Javeriana University campus, Bogota. 12th International Conference on Urban Drainage – ICUD 2011. Porto Alegre, Brazil.

Torres, A., S. Sandoval, K. Navarro and L. M. Pulido (2011). Strategies of maintenance in storm sewers in Bogota (Colombia) using estimates of operational and flooding reliabilities. 12th International Conference on Urban Drainage – ICUD 2011. Porto Alegre, Brazil.

Tran, D. H., A. W. M. Ng and B. J. C. Perera (2007). "Neural networks deterioration models for serviceability condition of buried stormwater pipes." *Engineering Applications of Artificial Intelligence* 20(8): 1144-1151.

Tran, D. H., B. J. C. Perera and A. W. M. Ng (2010). "Hydraulic Deterioration Models for Storm-Water Drainage Pipes: Ordered Probit versus Probabilistic Neural Network." *Journal of Computing in Civil Engineering* 24(2): 140-150.

Ugarelli, R., S. M. Kristensen, J. Rostum, S. Saegrov and V. Di Federico (2009). "Statistical analysis and definition of blockages-prediction formulae for the wastewater network of Oslo by evolutionary computing." *Water science and technology* 59(8): 1457-1470.

Ugarelli, R., G. Venkatesh, H. Brattebo, V. Di Federico and S. Saegrov (2010). "Historical analysis of blockages in wastewater pipelines in Oslo and diagnosis of causative pipeline characteristics." *Urban Water Journal* 7(6): 335-343.

UNAL (2010). Modelación Dinámica de la Calidad del Agua del Río Bogotá, Universidad Nacional de Colombia.

Uniandes (2001). Instrumentación y análisis ambiental de una subcuenca del sistema de alcantarillado de Bogotá - Informe Final, Universidad de los Andes.

Uniandes (2003). Bogotá River water quality modelling – Final Report.

Uniandes (2005). Definition and Technical Evaluation of Several Alternatives for Zoning, Operation and Managing of Muña Reservoir. Abstract for experts., Department of Civil and Environmental Engineering.

Vanrolleghem, P. A., L. Benedetti and J. Meirlaen (2005). "Modelling and real-time control of the integrated urban wastewater system." *Environmental modelling & software* 20(4): 427-442.

Vanrolleghem, P. A., C. Rosen, U. Zaher, J. Copp, L. Benedetti, E. Ayesa and U. Jeppsson (2005). "Continuity-based interfacing of models for wastewater systems described by Petersen matrices." *Water science and technology* 52(1/2): 493.

Verbanck, M. A., R. M. Ashley and A. Bachoc (1994). "International workshop on origin, occurrence and behavior of sediments in sewer systems - Summary of conclusions." *Water Research* 28(1): 187-194.

Vollertsen, J. and T. Hvitved-Jacobsen (2000). "Sewer quality modeling - a dry weather approach." *Urban Water* 2(4): 295-303.

Vollertsen, J., T. Hvitved-Jacobsen and A. H. Nielsen (2005). "Stochastic Modeling of chemical oxygen demand transformations in gravity sewers." *Water Environment Research* 77(4): 331-339.

Vollertsen, J., T. Hvitved-Jacobsen, Z. Ujang and S. A. Talib (2002). "Integrated design of sewers and wastewater treatment plants." *Water science and technology* 46(9): 11-20.

Vollertsen, J., A. H. Nielsen, W. Yang and T. Hvitved-Jacobsen (2005). "Effects of in-sewer processes: a stochastic model approach." *Water science and technology* 52(3): 171-180.

WHO-UNICEF (2004). Meeting the MDG drinking water and sanitation target: A mid-term assessment of progress, World Health Organization and United Nations Children's Fund.

WHO-UNICEF (2010). Estimates for the use of improved drinking-water sources: Colombia. Joint Monitoring Programme for Water Supply and Sanitation.

WHO-UNICEF (2010). Estimates for the use of improved sanitation facilities: Colombia. Joint Monitoring Programme for Water Supply and Sanitation.

Wilkie, P. J., G. Hatzimihalis, P. Koutoufides and M. A. Connor (1996). "The contribution of domestic sources to levels of key organic and inorganic pollutants in sewage: The case of Melbourne, Australia." *Water science and technology* 34(3-4): 63-70.

Willems, P. (2008). "Quantification and relative comparison of different types of uncertainties in sewer water quality modeling." *Water Research* 42(13): 3539-3551.

Willems, P. and J. Berlamont (2002). "Probabilistic emission and immission modelling: Case-study of the combined sewer- WWTP- Receiving water system at Dessel(Belgium)." *Water science and technology* 45(3): 117.

Wirahadikusumah, R., D. Abraham and T. Iseley (2001). "Challenging Issues in Modeling Deterioration of Combined Sewers." *Journal of Infrastructure Systems* 7(2): 77-84.

Zabel, T., I. Milne and G. McKay (2001). "Approaches adopted by the European Union and selected Member States for the control of urban pollution." *Urban Water* 3(1–2): 25-32.

Zheng, S. H. and V. Novotny (1991). "Stochastic modeling of combined sewer flows." *Water science and technology* 24(6): 35-40.