Booklets
Research & Development & Innovation

23

Canal de Isabel II

Observability study for hydraulic state estimation of the sectorised water supply network

Authors

Javier González Pérez

Roberto Mínguez Solana

Sarai Díaz García

Project Direction

Bruno Rodríguez Rodríguez

This document is an English translation of the original study published by Canal de Isabel II under the title:
*"**Estudio de Observabilidad para la estimación del estado hidráulico de la red sectorizada de abastecimiento**"*

Print ISSN: 2254-8955

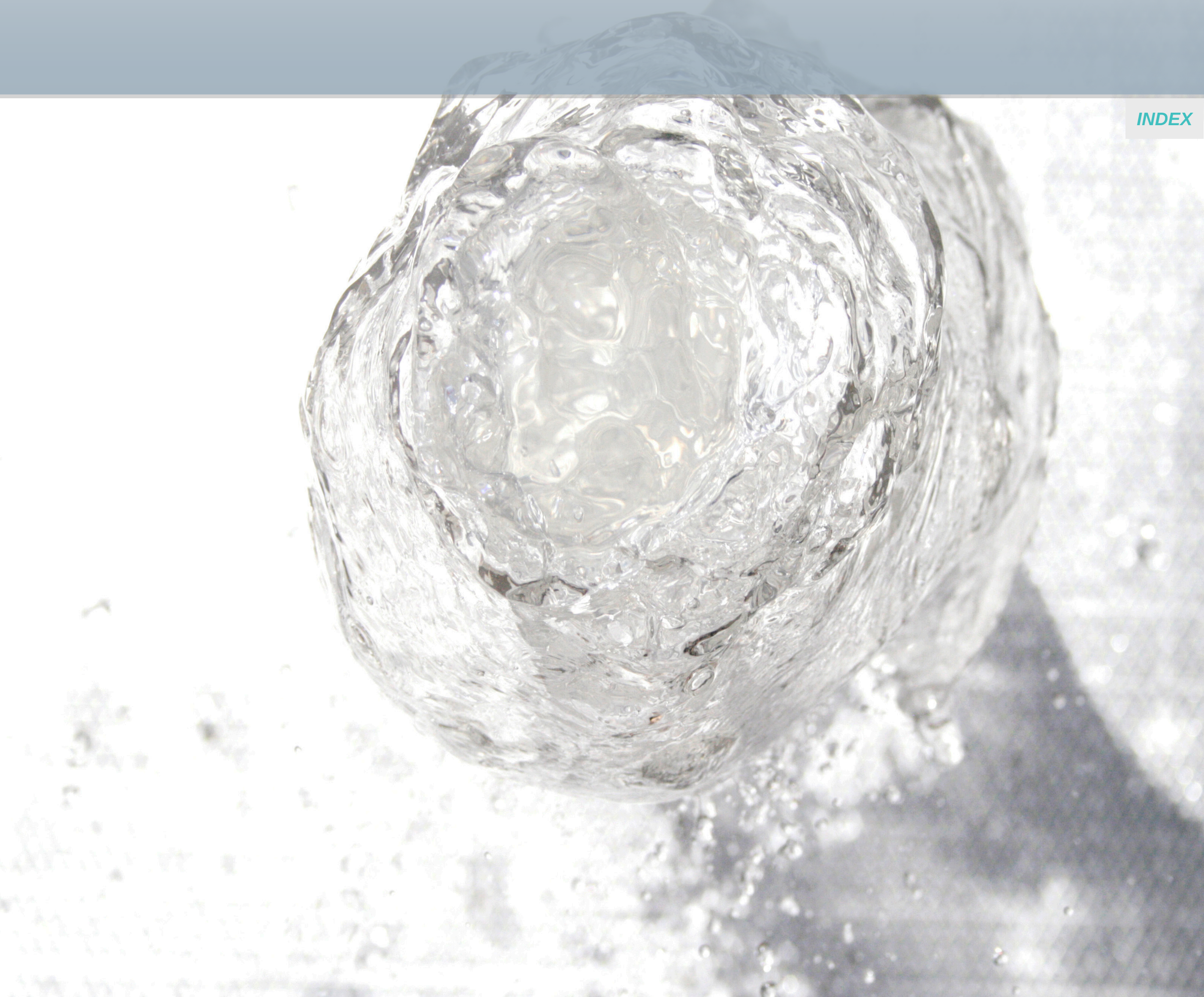
Electronic ISSN: 2340-1818

Legal Deposit (Spain): M-11431-2017

24

Observability study for hydraulic state estimation of the sectorised water supply network

INDEX



Liability Exclusion

The statements included in this document reflect the authors' opinions, and not necessarily those of Canal de Isabel II.

Neither Canal de Isabel II nor the authors of this document accept responsibility for any loss or damage that might be suffered by any institution or individual basing their actions on the contents of this document or the opinions expressed by the authors.

Presentation

The collection Booklets of Research, Development & Innovation of Canal de Isabel II are a part of the company's knowledge management strategy and of its R & D & I Plan.

These Booklets represent an element for diffusion of projects and initiatives developed and promoted by the company, and aim at innovation in areas related to the water services in an urban environment.

They deal with the problems tackled by each project as well as the results obtained. The aim of publishing these Booklets is to share experience and knowledge with the entire water industry sector, with the scientific community and with all those who work in the fields of research and innovation. With these publications what it is hoped is contribute the improvement and efficiency in water management and, as a result, make it possible to offer a better service to the citizens.

The titles published in the series to date are shown in the following table.

INDEX

BOOKLETS OF RESEARCH, DEVELOPMENT & INNOVATION PUBLISHED

Collection Number	***Research, Development and Innovation Booklets published***
1	Transferences of Water Rights between Urban and Agrarian Demands. The case of the Community of Madrid
2	Identification of Hydrometeorological Runs and Tendencies within the scope of the Canal de Isabel II system
3	Contribution of Canal de Isabel II to the International Demand Management Project (IDMF)
4	Microcomponents and Explanatory Factors on Domestic Water Consumption in the Comunidad de Madrid
5	Virtual Water and Hydrological footprint in the Comunidad de Madrid
6	Study on the saving potential of water for residential uses in the Comunidad de Madrid
7	Potentials of efficiency in using dishwashers in the Comunidad de Madrid
8	Accuracy in the measurement of individual water consumption in the Madrid Region
9	Research project to define and assess the applicability of a Bioassay Test to determine the toxicity of water using Zebra Fish embryos
10	Water Use Efficiency in Gardening in the Region of Comunidad de Madrid
11	Remote sensing techniques and geographical information systems for assessing water demand for outdoor uses in the Comunidad de Madrid
12	Cyanotoxin Dynamics Study in two of the Canal de Isabel II's supply reservoirs in the autonomous region of the Comunidad de Madrid
13	Development of a validation, estimation and prediction of hourly consumption by sector, for the distribution network of Canal de Isabel II
14	Monitoring of the consolidation urban development in the Comunidad de Madrid using remote sensing techniques
15	Experiences in the recovery of phosphorus from wastewater, in the form of Struvite, at Canal de Isabel II
16	Integration of weather forecasting in the management modules supply system of Canal de Isabel II, via daily contributions models
17	Improvement in forecast capacity of monthly and seasonal runoff in the scope of Canal de Isabel II
18	Inflow of nutrients from the basin to Pinilla reservoir. Effect on the eutrophication process
19	A new criterion for calculating urban sewage flows
20	Idea Management at Canal de Isabel II Gestión: The GENYAL Experience

Collection Number	***Research, Development and Innovation Booklets published***
21	Research on measuring techniques for subsidence related to groundwater exploitation
22	Precipitation patterns in the basins of the Lozoya and adjacent rivers

TABLE OF CONTENTS

1. Executive Summary

INDEX

Technical Data Sheet

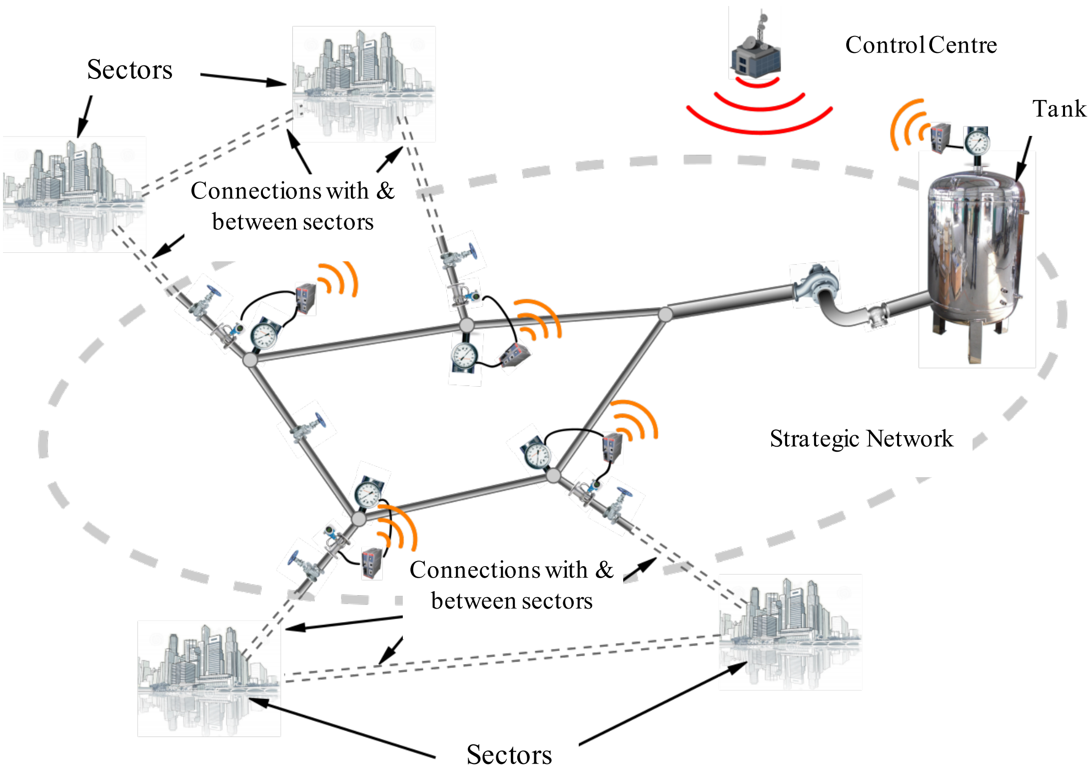
Project title	Observability study for hydraulic state estimation of the sectorised supply network
Research line	Ensuring strategic service continuity
Areas at Canal de Isabel II involved	Subdirección I+D+i
External participation	Hidralab Ingeniería y Desarrollo, S.L.
Aim and justification of the Project	Performing of the complete observability study (identification of observable islands and quantification of observation uncertainty) of a strategic supply network, as a step prior to the implementation of tools for real-time estimation of the hydraulic state of a system. Proposal of options to increase monitoring to improve the result of state estimation
Contribution to the state of the art	Adaptation to supply networks of techniques for identifying observable islands existing in the literature. Concept of topological observability. Large-scale implementation. Adaptation of techniques for the quantifying observation uncertainty to large networks, minimising numerical errors. Development of a strategy for locating additional metering devices to minimise system uncertainty.
Summary of the Project development and outstanding milestones	Review of the state of the art and the development of methodologies: adaptation to large networks. Identification of observable islands and observable conditions. Quantification of observation uncertainty. Proposal for strategic location for incorporating additional metering devices to improve uncertainty.
Summary of the results obtained	Observable islands under normal operating conditions and sensitivity to changes in network topology. Observation uncertainty in habitual demand scenarios. Strategic locations for new metering devices for different levels of investment.
Research Lines open for continuing the work	Development and implementation of techniques for real-time hydraulic state estimation of the sectorised supply network and derived applications.

Executive Summary

Canal de Isabel II, hereinafter Canal, has been making significant investments to upgrade and maintain its information technology systems, aimed at monitoring network operation, the final aim of which is to make information available on its state for operational decision making. Specifically, work has been performed to implement a SCADA[1] system to collect data in real time from the metering devices distributed throughout the supply network and to complement the corporate geographic information system and the existing calibrated hydraulic models for the sectorised strategic network or sub-divided into sectors. This network, known as strategic, is composed of the main arteries that supply water to the five hundred and eighty-two (582) sectors established in the Sectorisation Plan for the supply network, executed in the 2005-2015 period.

The SCADA system provides continuous, real-time information on the strategic network for specific locations, as shown schematically in Figure 1.

FIGURE 1. GENERIC DIAGRAM OF THE OPERATION OF A NETWORK SUBDIVIDED INTO SECTORS



[1] SCADA: Supervisory Control and Data Acquisition.

INDEX

This figure represents the strategic network, which is made up of the group of DWTPs (Drinking Water Treatment Plants), tanks, water distribution mains, pumps, valves, etc. and has as its purpose the transport of water from the sources (in the case of Figure 1, from the tank), to the output to the distribution sectors. The strategic network has a hydraulic model to characterise its behaviour, complemented by detailed models in the sectors, which are not the subject of this work. The metering devices installed send the corresponding measurements to the control centre in real time. The key to the system operating appropriately is for the large amounts of data provided by the monitoring system to be converted into improved operation and service to the user, with it being necessary to process all the information to transform it into real knowledge on the network, for it to be useful in decision making.

In this context, the use is proposed of so-called **state estimation techniques**, used successfully in other types of more highly instrumented distribution networks than transport water networks, such as electricity grids, as appropriate tools for that transformation of the isolated measurements into useful knowledge for managing and operating the system.

A "**state estimator**" is an algorithm that makes it possible, taking all the information collected by the SCADA system and the equations governing the hydraulic problem as its basis, to deduce the flow conditions at any time and at any point in the network area subjected to analysis. Therefore, its implementation can significantly improve water supply system operation, as it is a basic tool to provide detailed information about what is happening in the network or the future trends in it by processing thousands of signals, which it would be impossible to manage effectively in time manually. State estimation can be used to detect trends to be corrected, changes in system operation different to those planned or other types of incidents such as breaks, opening of sectors, errors in automatic metering sensors, etc. that could eventually cause interruptions or reduced service to the customer.

Assuming the schematic network in Figure 1 again, which shows some of the network elements, such as tanks, nodes, valves, pumps, supply sectors, together with some metering instrumentation, the purpose of state estimation is to:

1. Know the most probable value of the magnitudes defining the flow (flow rates and pressures) at any position within the network, based on the measurements and the hydraulic model of it, as well as calculating the uncertainty associated with this estimate.

2. Filter out possible errors associated with the metering instruments, when a device does not provide a measurement consistent with the rest of the measurements or consistent with the hydraulic state derived from them.

3. Detect possible associated inconsistencies and any possible incidents in the network that have not yet been notified, such as breaks, changes in points of water supply and consumption in the strategic network, operations in pumping stations, opening/closing valves or other elements capable of altering the hydraulic topology of the network.

This information is fundamental in large networks for controlling the system, as the volume of information to be considered prevents its interpretation by manual procedures, calling for effective automatic systems. The information that state estimation provides is key to understanding the situation of the hydraulic network and various tools can be derived from it to support network management and operation, such as tools for prognosis, planning, optimisation of operations, conserving equipment, etc.

However, before proposing a state estimation procedure, it is necessary to evaluate in what proportion the metering devices currently installed make it possible to observe what is occurring in it. This analysis in large networks in practice brings three challenges: i) performing it automatically, as although manual analysis would be possible, the size of the network to be analysed and the number of combinatorial analyses to be processed make it impracticable for manual procedures; ii) adapting the general formulation of the problem to the equations governing the hydraulic flow of water distribution networks, added to the peculiarity of dealing with networks for transporting water to distribution sectors; and iii) dealing with a problem which, due to its disparity of orders of magnitude in the dimension of its variables, is numerically ill conditioned and requires the use of special treatments.

The aim of this research work is to overcome these three challenges appearing in applying observability analysis techniques to the problem of flow in large water transport networks sub-divided into sectors. In a practical sense, its application is intended to tackle the following questions:

a. With the current configuration of measurements in the network, on which elements/variables of flow will it be possible to make an estimate of magnitude and which not? That is to say, which elements/variables of flow are in an observable island and which are not?

b. If the operation of certain elements that make up the network topology is put in doubt, such as pumps and valves, for which of these elements would it be possible to infer their most likely state and for which would there be insufficient information? This concept of topological observability is novel in the state of the art for water distribution networks and is derived from this study due to the practical problems that occur in operating large supply networks.

c. If a break occurs in the strategic network or any diversion not considered initially occurs, at which points in the network could their existence be detected and which not?

d. Given that all measurements are subject to errors and are measured with a given precision, there being no such thing as an exact measurement, the effect needs to be evaluated of the precisions of metering devices on the observation uncertainty in those elements/variables within an observable island of the network, responding to questions such as: What uncertainty does the estimate have? What minimum magnitude would a break need to have to be observable according to the metering devices available and the scenario of flow rates observed? Or, what alteration in flow does a change in topology need to produce to be identifiable?

As mentioned previously, this work is part of a more general challenge; that of applying state estimation as a tool to support improved management of the strategic network managed by Canal. To do this, a complete observability analysis was performed on that network, including the aforementioned processes of identification of observable islands and quantification of observation uncertainty. Finally, and a result of it having been possible to resolve the problem of observability in water transport networks, the problem of optimising the location of new metering devices to improve observability was tackled. The problem of optimisation this matter raises is defined in this work and an algorithm was developed to enable its resolution. This made it possible to provide a series of options as a result of this work to increase monitoring as a complement to the diagnosis of observability, by installing additional metering devices with the aim of improving the previous results and so also improving system response in a subsequent state estimation.

This project represents a contribution in itself, as never before has the observability of a large water supply system been analysed. Traditionally, the state of a network has been estimated on the basis of a supposed network topology and a forecast of demands at the points of consumption, which can be considered pseudo-measurements or measurements with an error level far above the automatic metering instruments. Up to now the number of metering points available automatically in a network was not high enough to attempt to understand its state based on the information from the remote metering system. However, the progressive increase in the number of metering points aimed at giving greater control over what happens in the network, a trend to which the sub-dividing of networks into sectors has contributed considerably, makes it possible to know not only the variables directly measured remotely, but also to **identify** a large set of **variables** (flow or pressure at the different points) and **conditions** (opening or closing of valves, pumping, alternative points of consumption or supply) in which the network may be. The observability study identifies those network situation variables/conditions for which the state estimation can give a response.

Outlined below are the fundamental parts and main contributions of this study.

STRATEGIC NETWORK MANAGED BY CANAL

The study was performed on the strategic network managed by Canal in the Comunidad de Madrid, which supplies the set of distribution sectors. Due to the dimensions of this system, prior knowledge of the network topology and of the metering devices existing in it is essential to identify the observable islands and quantify their uncertainty. In this respect, the strategic network that permits transport of drinking water in the Community of Madrid Region has of the order of 57,000 nodes and over 50,000 pipes, with very diverse flow conditions, in its 3,300 kilometres.

In turn, as regards the metering devices available, the network subject to study has a high level of instrumentation, with over 1,000 flow meters, almost 400 water meters, around 650 pressure sensors and 240 level gauges in tanks. A significant proportion of the devices is concentrated at the control of inlets to the sectors, as corresponds to the philosophy of sub-dividing a supply network into sectors. As in any system, faults can occur in these devices (sensor, telemetry, etc. faults), hence a part of the measurements existing is not available at certain times, which can lead to a reduction in the observable islands in the system or an increase in the uncertainty in them. In this regard, it should be emphasised that Canal has a tool of its own, **CHYPRE**, that provides estimates of consumption (pseudo-measurements) in the sectors that would be able to replace missing measurements in certain scenarios, although at the cost of a considerable increase in uncertainty. Observability makes it possible to know what variable/condition is identifiable in the network and with what precision or sensitivity threshold, making use of the set of information (measurements and pseudo-measurements) available in it at each time.

IDENTIFICATION OF OBSERVABLE ISLANDS

The identification of observable islands is an approach that makes it possible to evaluate whether it is possible to "**observe**" or estimate the various hydraulic variables (flow, demand and piezometric level) in a network zone for a given configuration of measurements. A variable is considered to be within an observable island when sufficient algebraic relationships exist to determine its value based on the measurements available. The basis for this technique consists of evaluating the number of linearly independent rows of the so-called Jacobian matrix for the system, which contains the basic relationships between the network variables. As such, the first contribution of this project is to adapt the construction of this matrix to a generic supply network, where the existence of singular elements specific to these systems (pumps, valves, consumption sectors, etc.) need to be incorporated. In this respect, the concept of "**topological observability**" is introduced in this project. This has not been covered in the scientific literature up to now. It consists of analysing whether it is possible to know about the state of the system pumps and valves as well as the hydraulic state of the network, in particular whether these elements are open or closed or, in general, the energy difference they produce between their inlet and outlet. On the other hand, the observability analysis was extended to supply sectors, in order to know whether it is possible to understand the way in which the sectors are supplied, as they have connections with various nodes in the strategic network and even between each other (see Figure 1). Unlike the strategic network, in the connections between sectors only the water balance equations are considered (not including the energy conservation equations), which requires adaptation of the analysis of observable islands. This adaptation is another of the contributions of this study, by including sector elements and connections between sectors for the first time, together with other types of elements which are already common in hydraulic networks, such as nodes, pipes, pumps, valves, tanks, etc., integrated here for the first time within a general framework for automatic identification of observable islands.

The next innovation consists of the adaptation of existing methodologies for the analysis of the Jacobian matrix, related to supply networks, which have mainly been developed for electrical systems, where the use of these techniques is the norm. Specifically, this work evaluated the performance of the so-called null space, optimisation, algebraic, binary and stochastic techniques, which were subjected to a multi-criteria analysis to identify the most suitable with a view to implementing it in the strategic network managed by Canal. After considering six criteria relating to the type of information involved in the process, its potential and computational cost, the algebraic method was found to be the most suitable, with its agility for combination with techniques for state estimation its most outstanding feature.

As stated previously, the analysis of observable islands is a dynamic procedure that depends heavily on the conditions under which it is carried out. However, given that the number of indoor valves, valves on the edge of sectors for opening or closing new points of supply and consumption or pump sets included in the network is very large and the combinatorial analysis of possible situations is unmanageable, efficient automatic analysis of them is necessary. To evaluate the ability to observe all the network topological control elements, they were evaluated one at a time, calling the condition of each element into doubt separately, also evaluating the possibilities of observing the state of the pumps and valves in case of a change of operation, the flows diverted to the distribution network from the strategic network and possible breaks in the main network.

QUANTIFICATION OF OBSERVATION UNCERTAINTY

Quantification of observation uncertainty is traditionally performed based on the Jacobian matrix for the system, which makes it possible to propagate the uncertainty of measurements to the variables using the FOSM[2] method, which is widely used in the hydraulic field and that in this case was applied to the unconstrained weighted least squares fit that the state estimation problem represents. Although the FOSM method is traditionally applied, its adaptation to the strategic network is not immediate, given the size of the network and the disparity of orders of magnitude in the variables that occur there. Specifically, significant differences in order of magnitude of flow conditions are seen, given that elements associated with practically nil velocities alternate with other elements subjected to intense flows, as well as ample differences in pipe sizes or hydraulic gradients. This variability required the adaptation of the traditional methodology to large networks in order to keep the numerical errors under control. Specifically, it was chosen to introduce constraints of a hydraulic type to the problem, also implementing a singular value decomposition (SVD[3]) technique to undertake a mathematical approach not discussed in the literature. With this methodology, the uncertainties associated with common scenarios in the strategic network were evaluated for maximum, mean and minimum demand scenarios.

STRATEGIC LOCATION OF ADDITIONAL METERING DEVICES

There is no doubt that the increase in system information on the basis of additional measurements has two immediate effects: the first is an increase in redundancy, which would make it possible to detect a larger number of inconsistencies, and the second is a reduction in uncertainty in the vicinity of the new device. Hence then, the identification of possible locations for installing additional metering devices can be performed in order to **extend the system's observable zones** in certain scenarios or to **reduce observation uncertainty**. Both approaches result in an improvement in system observability, with their effect on a subsequent state estimation process being notable. In this project, owing to the high level of instrumentation in the strategic network managed by Canal, where the diversions to the consumption sectors are monitored, the second option was chosen, with the intention of improving the quality of the observations.

The methodology used for locating additional metering devices represents a contribution in itself, as it is considered as an optimisation process seeking the most appropriate positions among the variables that recorded the highest values of uncertainty in the previous phase. The definition of the problem of optimisation, integrating the problem of identifying observable islands and estimating the associated uncertainty described above, contributes a very useful tool for investment planning. Specifically, work was performed with four levels of investment, making it possible to evaluate the sensitivity of overall system uncertainty to monitoring improvements. In this, a dimensionless parameter was chosen that evaluates the level of observability of each scenario with respect to a state of reference, making it possible to perform an evaluation of the improvement achieved for different ranges of investment, which can be evaluated on different types of variables in the network (consumptions, flows in pipes, pressures, etc.).

[2] FOSM: First Order Second Moment
[3] SVD: Singular Value Decomposition

Furthermore, the study shows that there are considerable differences within a single scenario between the precision index in the observation obtained on variables of the pressure type and variables of the consumption type, where the latter are observed with worse precision.

The observability study performed on Canal's network clearly shows the high potential that remote metering systems implemented in sectorised networks offer for understanding the state of the system in large transport networks, making it possible to extend the ability to observe and identify particular network conditions beyond the directly measured variables. It is shown, therefore, that hydraulic state estimation, as a subsequent dynamic tool, has a great potential capacity for monitoring the network state, giving early detection of significant incidents that have occurred and that have not been notified, which will make it possible to offer a high value service for decision making in operation.

2. Introduction

INDEX

2.1. STATEMENT OF THE PROBLEM

Canal de Isabel II manages the drinking water supply network for the Community of Madrid Region; one of largest water distribution networks in Europe, at over 17,000 kilometres in length. To do this it has sophisticated technologies and support systems in its management, which are continuously upgraded and modernised to improve the effectiveness and efficiency of network operation.

An example of one of these techniques is the **Plan for Sub-dividing the Supply Network into Sectors** which was executed over the 2005-2015 period, which involved identifying 582 sectors, also known as DMAs[4], which are zones isolated from each other hydraulically, but connected to the main arteries by means of a small number of inlet and outlet points. As such, the system structures a main network (called **strategic**) of approximately 3,300 kilometres supplying the existing sectors, which it is possible to connect to guarantee meeting demand in specific cases. This process represents a significant change in the way of understanding system management, making it possible to improve knowledge of the pattern of consumption and of the consequent circulation of water in the main network. Additionally, identifying the main arteries makes real-time monitoring of the network more viable, at least at a strategic level.

Precisely in this line, Canal has recently made significant investments relating to upgrading and maintaining information technology (ICT[5]) systems aimed at monitoring network behaviour and taking decisions consequently. Specifically, work has been carried out to implement a system for real-time collection of data from all the existing metering equipment (also known as SCADA), complementing the corporate geographic information system (GIS[6]) and the existing calibrated hydraulic models. However, in spite of this platform providing continuous, real-time information on the network, the large amount of data collected by the system does not automatically translate to improved operation and service to the user, rather there is a need to transform all the information into real knowledge of the network to be useful for decision-making, both in the short term (daily operations and bottleneck resolution) and in the medium and long term (expansion, maintenance, etc.). To do this, the most suitable techniques are those known as **state estimation techniques**.

In outline, state estimation techniques consist of implementing an algorithm that makes it possible, based on all the information collected by the SCADA system and the equations governing the hydraulic problem, to deduce the flow conditions at any time and at any point in the network. That is to say, state estimation is a technique that makes it possible to infer the hydraulic state of the system based on the set of data or measurements available. In turn, knowledge of the pressure at all the nodes and flows in all the pipes would lay the groundwork for subsequent evaluation of network operation in various forms (quality of service indicators, guarantee of supply, etc.) and in various scenarios (possible breaks, risk of supply cuts, etc.).

Hence the purpose of state estimation is to:

---

[4] DMA: District Metered Area
[5] ICT: Information & Communications Technologies
[6] GIS: Geographic Information System

1. Know the most probable value of the flow rates and pressures existing at any position within the network, based on the measurements and the hydraulic model of it, as well as to calculate the uncertainty associated with this estimate.

2. Filter out possible errors associated with the metering instruments, when a device does not provide a measurement consistent with the rest of the measurements or consistent with the hydraulic state derived from them.

3. Detect possible inconsistencies and any possible incidents in the network that have not yet been notified, such as breaks, changes in points of water supply and consumption in the strategic network, operations in pumping stations, opening/closing valves or other elements capable of altering the hydraulic topology of the network.

This information is fundamental in large networks for controlling the system, as the volume of information to be considered prevents its interpretation by manual procedures, calling for **effective automatic systems**. The information that state estimation provides is key to understanding the situation of the hydraulic network and various tools can be derived from it to support network management and operation, such as tools for prognosis, planning, optimisation of operations, conserving equipment, etc.

However, implementing state estimation techniques is not simple in practice. The reality is that not just any configuration of measurements is valid for characterising the state of flow and pressure in a certain network zone. For this reason, before proposing a state estimation procedure, it is necessary to evaluate the capacity of the current configuration of metering points in the network to observe what is happening in it. This is known as an **observability analysis**, which basically seeks to answer the following questions:

a. With the current configuration of measurements in the network, on which elements/variables of flow will it be possible to make an estimate of magnitude and which not? That is to say, which elements/variables of flow are observable and which are not? This process of identifying observable islands can also be performed for other reasons, such as to raising doubts on the operation of certain elements that make up the network topology (such as pumps and valves) or to evaluate the capability to detect breaks or the opening of diversions initially not considered.

b. Given that all measurements are subject to errors and are measured with a given precision, there being no such thing as an exact measurement, the effect needs to be evaluated of the precisions of metering devices on the observation uncertainty in those elements/variables that are observable, responding to questions such as: What uncertainty does the estimate have? What minimum magnitude would a break need to have to be observable according to the metering devices available and the scenario of flow rates observed? Or, what alteration in flow does a change in topology need to produce to be identifiable?

With the final aim of ensuring the proper implementation of a subsequent process for state estimation, this R&D&I project was executed with the objective of evaluating the applicability of observability measurement services to the network managed by Canal de Isabel II, quantifying, specifically for the target network, the level of observation the currently implemented automatic metering system provides. To do this, a methodology was proposed and researched and, finally, set out and used to perform a complete observability analysis on that network, including the aforementioned processes of identifying observable islands and quantifying observation uncertainty.

Also proposed in this work were a series of options to evaluate increased monitoring as a complement to the diagnosis of existing observability, by installing additional metering devices with the aim of improving the previous results and so also improving system response in a subsequent state estimation. The improvements in observation uncertainty were evaluated using a quantitative reference indicator.

Methodological Proposal

Due to all these reasons, the bulk of the work carried out in this study can be divided into four main activities, described by the methodological proposal to be performed in general terms:

- Analysis of the strategic network managed by Canal.
- Identification of observable islands under different assumptions.
- Quantification of the observation uncertainty under different assumptions.
- Proposal of locations for installing additional devices in the system in order to improve the previous results.

The carrying out of these four tasks made it possible to improve the state of knowledge in this field, prove the applicability of these techniques to large drinking water supply networks and assess the viability of implementing state estimation techniques in the strategic network managed by Canal. In this respect, it should be emphasised that this project represents a contribution in itself, as never before has the observability of a large water supply system been analysed. Traditionally, the state of a network has been estimated on the basis of a supposed network topology and a forecast of demands at the points of consumption, which can be considered pseudo-measurements or measurements with an error level far above the automatic metering instruments. Up to now, the number of metering points available automatically in a network was not high enough to try to understand its state based on the information from the remote metering system. However, the progressive increase in the number of metering points aimed at giving greater control over what happens in the network, a trend to which the sub-dividing of networks into sectors has contributed considerably, makes it possible to know not only the variables directly measured remotely, but also to identify a large set of variables (flow or pressure at the different points) and conditions (opening or closing of valves, pumping, alternative points of consumption or supply) in which the network may be. The observability study identifies those network situation variables/conditions for which the state estimation can give a response.

2.2. DOCUMENT STRUCTURE

Described briefly below are the various stages of the research project conducted, which form the structure of this document.

Formulation of the problem of state estimation and evaluation of the observability of the state: prior needs

The basic principles of the problem of state estimation in water networks are described. These show the need for performing the observability analysis of system state through the set of measurements available as a prior step.

Analysis of the strategic network managed by Canal de Isabel II

An analysis of the strategic network managed by Canal, as the subject of this study, is performed with the aim of identifying and characterising the network topology, simplifying the number of elements of which it is composed, for the purposes of the observability study.

Definition and processing of the hydraulic model equations governing the supply network

The network model and the process of constructing the Jacobian matrix, containing the relationships between the hydraulic variables and the existing measurements is described. These are essential to performing the observability analysis.

Identification of observable islands in the strategic network

The methodologies implemented at prototype scale to identify observable islands are presented, along with the results obtained under several assumptions with the methodology selected to be applied to the network managed by Canal, presenting the main statistics.

Quantification of the observation uncertainty in the strategic network

The methodology adopted in this research to quantify observation uncertainty is detailed, presenting the main statistics obtained from applying it under various assumptions.

Strategic location of additional metering devices

The methodology adopted and the results obtained as regards the location of additional metering devices in the strategic network based on the previous results are presented.

Conclusions

Finally, the conclusions drawn from the analysis undertaken are set out, with special emphasis on the viability of implementing state estimation techniques.

The document also includes a glossary of terms and acronyms used in performing this work.

INDEX

3. Formulation of the problem of state estimation and prior need for observability analysis of this state

INDEX

In general, a "**state estimator**" is an algorithm that needs to provide the most probable hydraulic state of the network given a set of measurements available at a given time, which will be provided by the SCADA system implemented. However, as revealed in the previous chapter, the results from state estimation are only reliable for the observable variables, which are those hydraulic variables (flows, demands and pressures) that can be deduced and observed based on the existing metering devices and the uncertainty of which can be quantified.

The basic aspects of formulating the problem of state estimation and the basis for the observability analysis is presented in context in the following sections, where the interest in installing additional metering devices to improve system operation is also made clear.

3.1. STATE ESTIMATION

The problem of state estimation is set out such that the existing set of measurements makes it possible to infer the so-called "*state variables"*, which constitute the minimum set of variables that enable the network to be characterised and based on which the rest of the system variables will be determined thanks to the equations governing the hydraulic problem.

As such then, in any supply network, a theoretical vector of measurements $\boldsymbol{z} \in \Re^m$ can be worked with, which could potentially include measures of pressure or consumption at the nodes, levels in the tanks or flow in the pipes. In turn, there will be a vector containing the state variables $\boldsymbol{x} \in \Re^n$, which throughout this research will be considered to be the piezometric levels at the nodes, as any combination of piezometric levels leads to a real solution for the network, which may not be true if the flows in the pipes are considered as state variables. Finally, there will be a non-linear relationship $\boldsymbol{g} : \Re^n \to \Re^m$ between measurements and state variables for a given system, resulting from the application of the mass and energy conservation equations.

This relationship can be written mathematically as:

$$\boldsymbol{z} = \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\varepsilon}\,, \tag{1}$$

Which represents a non-linear system of equations where $\boldsymbol{\varepsilon}$ are the errors associated with the measurements, which are typically assumed to be Gaussian with mean zero, i.e. $\mathrm{E}[\boldsymbol{\varepsilon}] = 0$ and variance-covariance matrix $\boldsymbol{R}_z$.

State estimation consists of finding the most probable values for the state variables by solving the following least squares problem:

$$\underset{x}{Min}\, F(\boldsymbol{x}) = \boldsymbol{\varepsilon}^T \boldsymbol{R}_z^{-1} \boldsymbol{\varepsilon} = [z - \boldsymbol{g}(\boldsymbol{x})]^T \boldsymbol{R}_z^{-1} [z - \boldsymbol{g}(\boldsymbol{x})], \tag{2}$$

Where $\hat{\boldsymbol{x}}$ is the optimum solution to the problem. Note that the errors are multiplied by the inverse of the variance-covariance matrix associated with the measurement errors, and as these are usually assumed to be independent, this is a diagonal matrix. The target function minimises the sum of the squared errors defined by the expression **(1),** making it possible to assign lower uncertainty to those measurements with errors with lower standard deviations.

The unconstrained problem, set out in the expression **(2)**, is usually solved using the normal equations method, which makes it possible to calculate the optimum solution for the state variables by solving the following system of linear equations iteratively:

$$[\boldsymbol{J}_{(\nu)}^T \boldsymbol{R}_z^{-1} \boldsymbol{J}_{(\nu)}]\Delta\hat{\boldsymbol{x}}_{(\nu+1)} = [\boldsymbol{J}_{(\nu)}^T \boldsymbol{R}_z^{-1}](\boldsymbol{z} - \boldsymbol{g}(\hat{\boldsymbol{x}}_{(\nu)})), \quad (3)$$

Where $\boldsymbol{J}_{(\nu)} \in \Re^{mxn}$ is the measurement Jacobian matrix (see point 5.3. "Construction of the Jacobian matrix") at the optimum $\hat{\boldsymbol{x}}_{(\nu)}$ and ν is an iteration counter. The optimum solution can be progressively updated as $\hat{\boldsymbol{x}}_{(\nu+1)} = \hat{\boldsymbol{x}}_{(\nu)} + \Delta\hat{\boldsymbol{x}}_{(\nu+1)}$.

3.2. OBSERVABILITY ANALYSIS

According to the equation **(3)**, a theoretical condition sufficient for the existence of a unique solution to the problem of estimation defined by the expression **(2)** is for the system to be determinate compatible, i.e. that the matrix $\boldsymbol{J}$ is full rank n. If the Jacobian matrix is full rank, the matrix $[\boldsymbol{J}_{(\nu)}^T \boldsymbol{R}_z^{-1} \boldsymbol{J}_{(\nu)}]$ can be inverted and the system is observable, but apart from this reflection, observability has received little attention in the study of water distribution networks. The following section presents how this first aspect of observability analysis can be tackled, which basically evaluates the algebraic relationships between variables to identify those that can be estimated based on the existing measurement configuration. The challenge in large networks is to make these checks automatically and efficiently, as it is always implemented in very simplified examples in the literature.

However, to analyse the observability of hydraulic variables it is not sufficient to evaluate the algebraic relationships existing between them, but it is also necessary to quantify the uncertainty with which it can be hoped to know the variable thanks to the available metering devices. In point 3.2.2. "Quantification of observation uncertainty", the importance of this quantification in large supply systems is explained, as well as presenting the methodology traditionally adopted for this. Subsequently, within this record, the methodological proposal resulting from this work to tackle the problem of large networks is set out.

3.2.1. Identification of observable islands

As stated at the start of this section, a first aspect to be tackled for observability analysis is the identification of observable islands, which consists of detecting which variables can be deduced based on the existing measurements in the network thanks to the equations governing flow. To do this, it suffices to evaluate the algebraic relationships existing between the variables thanks to the aforementioned equations, which constitute the hydraulic model of the network. Although this model and the construction of the associated Jacobian matrix will be presented in detail in the chapter entitled "**Modelling the supply network for observability analysis**", the idea is that it will be necessary to assess the relationships between variables. As stated previously, the basic idea is to evaluate the rank of the matrix, i.e. to identify the number of linearly independent rows or columns with the aim of identifying what state variables could be deduced on solving the problem of state estimation set out in equation **(3).**

One of the reasons for the low level of implementation of techniques for the identification of observable islands in water supply networks is the fact that once the levels in the tanks and the consumption in the nodes are known, the flow can always be solved or, to put it another way, the network is always observable from the point of view of having sufficient algebraic relations to determine all the system variables. However, its solution depends on assuming hypotheses on the value of real consumption, network operating conditions (roughnesses, local energy losses, among others), etc. Such hypotheses can have a high level of uncertainty. In this regard, if, to determine the state of a supply network, there is baseline information with the consumption produced at each point and levels in tanks and neither the parameters for the various elements nor the topological characteristics of the network are in doubt, the network is observable from the mathematical point of view (the set of network elements are in an observable island) even in the absence of any additional measurements, although in practice the uncertainties linked to this observation can be very high.

In spite of this, if it is wished to control and limit the uncertainty linked to network operation, measurements need to be introduced of variables which, even with these variables being previously observable, reduce the uncertainty in their estimated value and also help to reduce the uncertainty in the rest of the variables not measured directly, before relying on the representative nature and calibration of the model of hydraulic operation. In this case the problem is handled as a problem of propagation and quantification of uncertainty, without doubting the existence of sufficient algebraic relations to keep system observable.

However, when there are pumps and valves in a network and the variability of its state and operation are in doubt, these elements introduce unknowns to the system that make it necessary to resort to additional measurements or assumptions to ensure that these elements are observable. It is not possible to know whether a pump or a valve is open or closed without sufficient measurements, either direct or indirect, related to its state of operation. In this case, from the point of view of evaluating the availability of sufficient algebraic relationships, observability analysis is essential as a starting point to discover if the state of these elements is observable. Once this condition has been met, the possibility of actually observing their state in practice will also depend on the uncertainty linked to the measurements and the derived variables dependent on or related to these measures.

On the other hand, the operation of metering devices is not perfect and sporadically some of them stop transmitting information due to a fault of some sort, either in the device itself or in the transmission procedure. This means the configuration of measurements available can vary over time. Hence it is necessary to develop an efficient, automatic procedure to make it possible to deduce the elements within an observable island for a given measurement configuration and to be able to evaluate it dynamically.

It is important to emphasise that the Jacobian matrix plays a crucial role in the analysis of the observable islands in the system, as it gathers the basic relationships between variables. Additionally, the matrix maintains the structural relationships between measurements and state variables even if the expression **(1)** is linearised with respect to any point $\boldsymbol{x_0}$:

$$\Delta z = \boldsymbol{J}_0 \Delta \boldsymbol{x} + \Delta \varepsilon \text{ ,} \qquad \text{(4)}$$

Where $\Delta z = z - g(\boldsymbol{x_0})$ is the residual measurement vector, $\Delta \boldsymbol{x}$ is the incremental variation of the system state and $\Delta \varepsilon$ corresponds to the incremental variation of the errors. The matrix $\boldsymbol{J}$ is now accompanied by a suffix ($\boldsymbol{J}_0$) to indicate that it needs to be particularised for a real possible physical state $\boldsymbol{x}_0$, i.e. that observability is analysed for specific flow scenarios. Hence, once the scenario which is the subject of the scenario has been established, identifying observable zones requires implementation of a mathematical technique that makes it possible to evaluate the rank of the Jacobian matrix. In the chapter entitled “Identification of observable islands in the strategic network”, the possible strategies are dealt with in detail.

3.2.2. Quantification of observation uncertainty

In addition to evaluating the potentially observable zone according to the existing relationships between variables, the precision with which these could be observed needs to be judged. In this respect, it should be mentioned that state estimation and observability analysis techniques were conceived in the 1970s with the aim of characterising the state of large electrical systems, which are typically well instrumented and have redundancy ratios (the ratio between the number of measurements existing and the minimum number of measurements required to achieve observability) far in excess of one. Water supply networks, though, have traditionally been characterised by their low level of instrumentation. To make good this scarcity, the lack of real measurements has been complemented with predictions of consumption, also known as *pseudo-measurements*. In this regard, it should be said that all measurements, either from devices or predictions, are associated with an uncertainty which causes a deviation in the state estimation, with the uncertainty being considerably higher when it involves *pseudo-measurements*, as these are only estimates based on previously recorded data. For this reason, quantification of observation uncertainty is fundamental in water supply systems to distinguish the improvement that a set of measurements produces in precision in comparison to a situation where only *pseudo-measurements* are available.

In the specific case of the strategic network managed by Canal, the process of sub-dividing it into sectors has enabled better monitoring of consumptions existing in the Community of Madrid Region, with flow rate metering devices located at the outlets to the majority of the sectors during the transformation. In parallel, Canal has developed a computer application called CHYPRE[7], which makes it possible to estimate sector consumption and its associated uncertainty. The information provided by CHYPRE is, after all, *pseudo-measurement* of consumptions in each of the sectors, which could replace the corresponding flow metering device reading in case of failure of the latter for various reasons (sensor fault, telemetry fault, etc.), with the consequent increase in system uncertainty.

Observation uncertainty has been traditionally quantified by using the expressions used for evaluating the uncertainty of a weighted least squares fit, like the state estimation problem. To do this, the linearised relationship between state variables and measurements is used:

$$\Delta\hat{\mathbf{x}} = [\mathbf{J}^T \mathbf{R}_z^{-1} \mathbf{J}\ ]^{-1} [\mathbf{J}^T \mathbf{R}_z^{-1}]\ \Delta\mathbf{z} \tag{5}$$

Where in this case $\Delta\mathbf{z} = \mathbf{z} - \mathbf{g}(\hat{\mathbf{x}})$ and $\mathbf{J} \in \Re^{mxn}$ is the measurement Jacobian matrix evaluated in the optimal solution of the problem **(2)**.

In this way, by applying the FOSM[8] method to this expression to propagate the uncertainty in the available measurements to the state variables, the variance-covariance matrix can be obtained for the state variables $\mathbf{R}_{\hat{x}}$, which is:

$$\mathbf{R}_{\hat{x}} = [\mathbf{J}^T \mathbf{R}_z^{-1} \mathbf{J}\ ]^{-1} \tag{6}$$

Note that a sufficient theoretical condition for quantifying the uncertainty is that the matrix $\mathbf{J}$ is full rank, i.e. that the system is observable. This condition implies that the inverse can be calculated, leading to the obtaining of the variance-covariance matrix in **(6)**.

Once $\mathbf{R}_{\hat{x}}$ has been calculated, the variance-covariance matrix for the rest of the rest of hydraulic variables $\mathbf{R}_{Q,q}$ (either flows Q or demands q) can be deduced by applying the FOSM method again:

$$\mathbf{R}_{Q,q} = \mathbf{J}_{Q,q} \mathbf{R}_{\hat{x}} \mathbf{J}_{Q,q}^{\ T} \tag{7}$$

Where $\mathbf{J}_{Q,q}$ refers to the part of the measurement Jacobian matrix that relates the flows and the demands to piezometric levels (taken as state variables in this project), respectively.

[7] CHYPRE: AN APPLICATION THAT ENABLES ESTIMATION OF CONSUMPTION IN SECTORS AND THE ASSOCIATED UNCERTAINTY IN THE NETWORK MANAGED BY CANAL GESTIÓN. SEE CARRASCO AND GARCÍA (2011) FOR FURTHER DETAIL.

[8] FOSM: FIRST ORDER SECOND MOMENT

In this way, the observation uncertainty is obtained for all the system hydraulic variables, with the aim in the subsequent phase of locating additional metering devices to minimise its value, as described briefly below.

3.3. STRATEGIC LOCATION OF ADDITIONAL METERING DEVICES

The problem commonly called "**optimal placement of metering devices**" refers to the approach to enable objective selection of the most suitable locations to install additional metering devices in order to improve system performance with respect to a given criterion. In this particular study, the location of additional metering devices will be referred to at all times with the aim of improving system response for state estimation, which is the ultimate end at which this study is aimed, conditioned at all times by a given level of investment.

To this end, the methods for locating additional metering devices can be classified into two major groups:

- **Methods based on ensuring minimum conditions of observability:** In this case the aim is to increase the resilience of the metering system by guaranteeing the observability of certain variables or the network as a whole in the case of the loss of a measurement, without the uncertainty of the estimate mattering as a first approximation, although this is related to the number of metering devices installed.

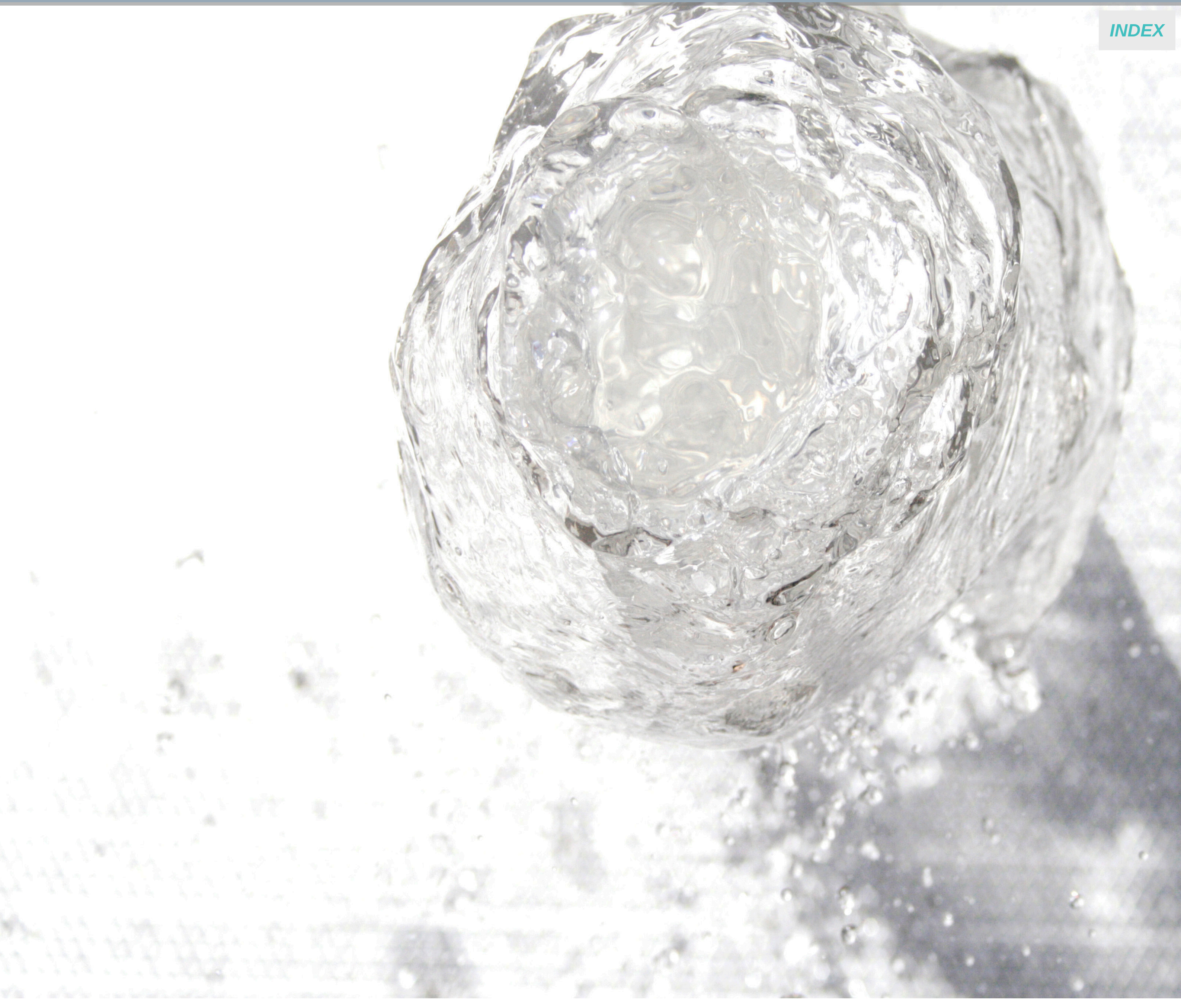
- **Methods based on reducing the uncertainty:** In this case the aim is to place new devices so that they reduce the uncertainty of the variables located in observable islands, which can contain the set of system variables and elements, so as to achieve greater certainty on network operation, facilitate its management, enable detection of breaks, etc.

In any case, adding new metering devices generally improves both conditions, but the difference between these two options lies in the definition of the objective function of the optimisation problem into which each of the strategies is translated. The option selected in this work was the latter, because, as will be seen in the chapters entitled "**Identification of observable islands in the strategic network**" and "**Quantification of the observation uncertainty in the strategic network**", the fact of having measurement estimates from the CHYPRE system and information with differing levels of precision on the levels in tanks produces situations in which the set of elements of which the network is made up covers the entire network, whereas it gives uncertainties that can be improved in the observation of the various system hydraulic variables.

In this way, the strategy of locating new devices will attempt to reduce the uncertainty in the most relevant system variables, which will indirectly improve system resilience to the loss of any measurements. The methodology implemented in the strategic network with this aim will be explained in detail in the chapter "**Strategic location of additional metering devices**", as it was largely determined by the result of the observability and uncertainty analysis conducted.

4. Analysis of the strategic network

INDEX



This chapter covers the analysis of the strategic network managed by Canal and the set of metering devices available in it. This puts the real situation on which this study was conducted into context.

4.1. THE NETWORK

The strategic network managed by Canal transports drinking water from the treatment stations to the set of sectors that make up the distribution network. In outline, the elements that make up the strategic network are those described in Table 1.

TABLE 1. ELEMENTS THAT MAKE UP THE STRATEGIC NETWORK

Elements	*Number of elements*
Tanks	310
Nodes without demand	56,514
Nodes with demand	697
Pipes	44,333
Shut-off valves	14,344
Regulating valves	602
Pumps	447

4.2. ANALYSIS OF METERING DEVICES

Once the network has been characterised, the other essential element that makes it possible to identify the observable variables and quantify observation uncertainty is the set of measurements available, which determines the success of state estimation. Identified in this section are the metering devices in the system, linking these to one or more hydraulic variables within the network model, which will be described in detail in the next chapter.

4.2.1. Metering devices available

There are three major types of metering devices in the network:

- Flow meters and water meters that make it possible to measure flow rate. Table 2 shows the statistics for these devices, where it can be seen that 55% of the flow meters and 48% of the water meters located on the strategic network are used to measure flow in pipes, while the rest makes it possible to monitor flows diverted at the inlet to the sectors.
- Pressure meters. There are 490 pressure sensors located alongside flow meters and 170 pressure sensors next to water meters. A further 40 pressure loggers are also located in the network, the height of which will also need to be known for a subsequent state estimation process.
- Level gauges in tanks. The total number of tanks with at least one metering device is 241.

TABLE 2. FLOW METERS AND WATER METERS IN THE STRATEGIC NETWORK

	Classification	Number of elements
Flow Meters	Strategic network total	1,112
	Measuring flow in network pipes	609
	Measuring flow at the inlet to sectors	503
Meters	Strategic network total	399
	Measuring flow in network pipes	190
	Measuring flow at the inlet to sectors	209

In this regard, there are other remotely measured signals, such as the pressure at other points in the network, the position of valves or the status of pumps, etc. However, at the time of performing this study it was not possible to identify reliably the variable or network state that the majority of them measure specifically, hence they were discounted in this phase, with the number of devices installed indicated in Table 2, being larger in practice.

4.2.2. Measurement precision

The precision of the measurements in the network determines the quality of the results of the observability analysis. Specifically, in this research it is essential for quantifying the observation uncertainty.

Among the factors that can have a major impact on measurement precision is the minimum measurement resolution, the measurement principle and other errors, such as those due to the lack of periodic calibration and general deterioration of the device due to age or level of use. The measurement precisions shown in Table 3 were assumed in this work.

TABLE 3. METERING DEVICE PRECISION

Metering device	Classification	Number of elements
Flow meters and water meters	Flow rate if speed > 0.3 m/s	2% + RIE*
	Flow rate if speed < 0.3 m/s	5% + RIE*
Pressure meters	Pressure at a node	0.025 bar
Level gauge	Tank level	0.025 m

**RIE: Resolution-Induced Error, considered to be 1 m^3 with hourly readings.*

4.3. JUSTIFICATION OF THE NEED FOR OBSERVABILITY ANALYSIS

As stated under the “Statement of the problem” point, the **observability analysis** and in particular the **identification of observable islands**, have received little attention in water supply systems. This is due to the fact that water distribution systems are usually considered to be entirely within an observable island given that:

1) estimates or predictions of demand are adopted based on historical logs to counteract the scarcity of metering equipment existing (*pseudo-measurements*); and

2) the state of the elements that can alter network topology, such as pumps and valves, is usually assumed to be known.

However, these assumptions are not valid for the case of the strategic network managed by Canal for the reasons stated below.

On the one hand, the strategic network has its own system called CHYPRE that provides real-time pseudo-measurements or estimates of consumption in the sectors, as well as their associated uncertainty. However, these estimates are associated with uncertainty values that are much higher than those of the metering devices available in the network, which offer greater certainty with a view to the subsequent implementation of a state estimation procedure. Nevertheless, and in spite of the large number of devices installed in the system, not all the devices provide measurements in real time, as in any distribution network an isolated fault of a device can occur, whether due to a sensor fault, a fault in the telemetry system, etc. This loss of information can lead to a deterioration of the observable islands existing in the network or an increase in the uncertainty of certain system variables. For example, if a fault occurs in a flow meter measuring the outlet to a sector, it may be possible to replace this with a pseudo-measurement (provided by CHYPRE), leading only to an increase in state estimation uncertainty. In turn, if a pressure sensor fails and there is no alternative redundant measurement, this fault could lead to this variable no longer being observable. Note that in either of the two cases a deterioration in the state estimation result would occur, either because the uncertainty increases or because certain variables cease to be observable.

Also needing to be added to this reality is the large number of exceptional elements in the strategic network. In this regard, the existence of a pump or a valve adds and unknown to the system, as its state can be open or closed depending on network operation and/or flow conditions. In this way, even when the normal operating conditions are known in the strategic network, situations can arise in which deviations occur with respect to this situation due to an operation, a mechanical fault, etc. in any of the nearly 450 pumps and 15,000 valves. If alterations of this kind occur, which invalidate the characteristic curves assumed, not all the variables are necessarily within an observable island, hence a tool is needed to make it possible to identify the observable elements and/or conditions in different situations, as well as the uncertainty of the estimate that would be obtained in each of them.

Finally, it needs to be borne in mind that one of the reasons for implementing SCADA systems to monitor water networks is to detect leaks. As such, the presence of a break from the modelling point of view can be regarded as analogous to the presence of a demand. This means that an additional unknown is added to the system and it is therefore necessary to know if the presence of a break in a specific location could be detected by the neighbouring instruments. In this specific case, the presence of breaks at transit nodes was analysed, such that the demand goes from nil to an unknown demand if there is a leak.

Therefore, the possibility of faults in existing metering devices and the large number of exceptional elements that can alter the network topology and the possibility of breaks existing make it necessary to carry out a full observability study as a step prior to state estimation. As such, the identification of observable zones and the quantification of estimation uncertainty in various scenarios will make it possible to make the most of the investments made in monitoring.

5. Definition and processing of the hydraulic model equations governing the supply network for observability analysis

INDEX

In view of the need to conduct the observability analysis, as a step before state estimation in the strategic network managed by Canal, it is necessary to define how supply network modelling is to be performed. According to the proposed network model, also presented in this chapter is the process for constructing the Jacobian matrix for the system, which is essential for implementing the procedures to identify observable islands and quantifying the observation uncertainty in this research, as stated in the chapter “Formulation of the problem of state estimation and prior need for observability analysis of this state”.

5.1. BASIC COMPONENTS OF THE SUPPLY NETWORK

Any water supply system can be represented as a network $N=(V,L)$, made up of a set of vertices or nodes (V) connected to each other by means of a set of pipes or piping (L). In turn, the following can be distinguished within the category of vertices or nodes:

- Demand or input nodes V^Q, where water is extracted from or introduced to the system, respectively
- Transit nodes V^T, where water does not leave or enter the system, only continues
- Tanks V^D, in which water is stored

The distinction between these is important with a view to assembling the hydraulic model.

As regards the pipes, a signing criterion is used according to which the flow direction in the pipes is assumed positive when the water moves from a lower numbered node to a higher numbered one. Consequently, two sets Ω_i^O and Ω_i^I can be defined for each node i, corresponding to output flows from node i to the rest of the nodes with numbering $j>i$ connected with i through a pipe and input flows to node i from the rest of the nodes with numbering $j<i$ connected with i through a pipe, respectively.

5.2. THE NETWORK MODEL

Once the basic components of the hydraulic model of the network have been defined, the variables associated with each of them, the possible measurements in the network and the relationships existing between the different variables that constitute it need to be established. It is important to highlight that this work assumes a pseudo-static approach that considers flow as permanent, so that successive times can be analysed as if they were independent. For this reason, the variables and the equations involved are considered independent of time t.

5.2.1. Variables to be considered

Starting with the definition of the variables, in general, the hydraulic variables involved in a generic water distribution network at a given time are:

- Pressure at each node ($p_i;\forall i \in V$). A network has as many pressure variables as there are nodes $n = n_n + n_d$ in the network, where n_n represents the number of nodes that are not tanks and n_d the number of tanks.
- Piezometric level at each node ($h_i;\forall i \in V$). As in the case of pressure, there are $n = n_n + n_d$ piezometric level variables.
- Flow in each pipe ($Q_{ij};\forall ij \in L$). There are as many variables for flow rate in the pipes as there are pipes n_p in the network.
- Demand (consumption) or input to the system at each node ($q_i;\forall i \in (V^Q \cup V^T)$), which is positive for inputs, negative for demands and zero in the transit nodes. There are as many variables of demand as there are nodes that are not tanks n_n in the network. Note that the existence of consumptions is only considered at nodes that are not tanks, as the tanks act as network boundary conditions, for which reason they will only be taken into account in terms of piezometric level in the network model adopted.

The rest of the parameters needed to define the system state are:

- as the height of the nodes ($e_i;\forall i \in V$),
- the length ($L_{ij};\forall ij \in L$),
- diameter ($D_{ij};\forall ij \in L$) and
- roughness of the pipes ($r_{ij};\forall ij \in L$),

They are assumed to be known for the problems of observability analysis and state estimation presented here as it is understood that these have already been the subject of a previous calibration. Note also that pressure and piezometric level are the same variable, as one can easily be converted to the other considering the height of the node $h_i = p_i + e_i$. In turn, and as stated in the chapter “Formulation of the problem of state estimation: prior needs”, the state variables are precisely considered to be the piezometric levels at the nodes.

5.2.2. Possible measurements

There are as many possible measurements as the model contains. Therefore, there are three types of possible measurements, one for each of the types of variables involved:

- Piezometric level at each node ($\tilde{h}_i ; \forall i \in V$)
- Flow in each pipe ($\tilde{Q}_{ij} ; \forall ij \in L$)
- Demand or consumption at each node that is not a tank ($\tilde{q}_i ; \forall i \in (V^Q \cup V^T)$)

The **~** indicates that these are measurements, which can be either metering device readings or pseudo-measurements, with the only difference being the associated uncertainty with a view to subsequent state estimation.

In this way, the following vector includes all possible measurements in the network:

$$z = (\tilde{h}_i ; \forall i \in V, \tilde{Q}_{ij} ; \forall ij \in L, \tilde{q}_i ; \forall i \in (V^Q \cup V^T))^T \quad \textbf{(8)}$$

5.2.3. Equations governing hydraulic behaviour

The relationships between the different variables involved, or to put it another way, between the existing measurements and the system variables, constitute the hydraulic model of the network. These relationships are defined below, expressed as a function of the state variables by means of the mass conservation or balance equation at the nodes that are not tanks ($\forall i \notin V^D$) and of the energy conservation or loss equation along the pipes:

- Piezometric level at each node $\tilde{h}_i$: if any piezometric level is measured, the corresponding state variable is immediately known, i.e. $\tilde{h}_i = h_i ; \forall i \in V$.
- Flow in each pipe $\tilde{Q}_{ij}$: the measurement of any flow in a pipe can be related to the piezometric levels at its end nodes i and j according to the following expression:

$$\tilde{Q}_{ij} = \frac{1}{K_{ij}^{\frac{1}{b}}} (h_i - h_j) \,|\, h_i - h_j \,|^{\frac{1}{b}-1} ; \forall ij \in L , \quad \textbf{(9)}$$

Where b is an exponential coefficient that depends on the loss equation considered, that is, $b = 1.852$ for the Hazen-Williams equation or $b = 2$ for the Darcy-Weisbach or Manning equation and K_{ij} is the roughness coefficient for the pipe. Note that this works with absolute values to enter the effect of the sign.

- Demand at each node $\tilde{q}_i$: the measurement of any demand $q_i; \forall i \in (V^Q \cup V^T)$, makes it possible to establish a balance with the input and output flows to and from the node in question using the mass conservation equation, which for incompressible fluids is:

$$\tilde{q}_i = -\sum_{\forall j \in \Omega_i^I} Q_{ij} + \sum_{\forall j \in \Omega_i^O} Q_{ij}; \forall i \in (V^Q \cup V^T), \quad \textbf{(10)}$$

Where $\tilde{q}_i$ is considered negative if the water leaves the system, positive if it is an input or zero if it is a transit node ($\tilde{q}_i = 0; \forall i \in V^T$). In addition, it is possible to substitute the flows with the piezometric level at the nodes using the expression **(9)**, with the demands then expressed as a function of the state variables.

5.3. CONSTRUCTION OF THE JACOBIAN MATRIX

5.3.1. Basic Jacobian matrix

The measurement Jacobian matrix for the system includes the first-order derivatives for all the variables that can be measured in the network with respect to piezometric levels at the nodes, which are the state variables. As such, the Jacobian matrix has as many columns as nodes n, each of which is associated with a piezometric level $h_i; \forall i \in V$, and as many rows are the total number of measurements m that can be taken in the system, represented by the vector given by the equation **(8)**.

The structure of a generic Jacobian matrix for any water distribution network is:

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial \tilde{h}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{h}_n}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_n} \\ \frac{\partial \tilde{Q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{Q}_{n_p}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_n} \\ \frac{\partial \tilde{q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{q}_{n_n}}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_{n_n}}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_{n_n}}{\partial h_n} \end{bmatrix} \quad \textbf{(11)}$$

Where n_p and n_n represent, respectively, the number of pipes in which the flow can be measured (total number of pipes) and the number of nodes where the demands can be measured and/or estimated (total number of nodes that are not tanks).

In this way, the first block of the Jacobian matrix corresponds to the identity matrix I_n with dimensions $n \times n$, as if the piezometric level is measured at a node, the variable is observable at that point.

In turn, the second block, referring to the flow in the pipes, can be calculated as:

$$\frac{\partial \widetilde{Q}_{ij}}{\partial h_k} = \begin{cases} \dfrac{1}{bK_{ij}^{\frac{1}{b}}} | h_i - h_j |^{\frac{1}{b}-1} \; si \; k = i \\ -\dfrac{1}{bK_{ij}^{\frac{1}{b}}} | h_i - h_j |^{\frac{1}{b}-1} \; si \; k = j; \; \forall ij \in L, \\ 0 \; in \; other \, case \end{cases} \tag{12}$$

Where i and j ($i < j$) correspond, respectively, to the pipe start and end node.

Finally, the third block is associated with demand measurements, the partial derivatives for which can be expressed as:

$$\frac{\partial \widetilde{q}_i}{\partial h_k} = -\sum_{\forall j \in \Omega_i^I} \frac{\partial Q_{ij}}{\partial h_k} + \sum_{\forall j \in \Omega_i^O} \frac{\partial Q_{ij}}{\partial h_k}; \forall i \in (V^Q \cup V^T), \tag{13}$$

Where the partial derivatives of the flows with respect to the piezometric levels involved can be obtained using the expression **(12).**

Once the structure of the Jacobian matrix has been understood, it is important to remember that this needs to be calculated for a given state of flow, so that the matrix **(11)** is particularised ($\boldsymbol{J}_0$) for a possible real physical state $\boldsymbol{x}_0$. Furthermore, the hydraulic state adopted must comply with additional condition of not having null flow in any of the pipes, as this particularity would lead to a mathematical indeterminacy $\frac{1}{0}$ in the expressions $\frac{\partial \widetilde{Q}_{ij}}{\partial h_k}$, producing a numerical ill-conditioning of the Jacobian matrix. Finally, to apply it to identify observable islands, the normalised matrix of $\boldsymbol{J}_0$ can be used, which is calculated by dividing the elements of each of them by the maximum absolute value of each row. This strategy reduces the numerical errors deriving from the application of the different methods for observability analysis.

5.3.2. Inclusion of pumps and valves

In view of the analysis of the strategic network carried out in the previous chapter, the importance is noted of elements of the pump and valve type in the network managed by Canal and, in general, for the operation of any network. The presence of these elements gives flexibility to network operation, enabling certain pipes to be opened/closed and the pressure regime in the system adjusted. Therefore, although the state of the pumps and valves is known beforehand under normal operating conditions, this can cease to be the case in some circumstances as the result of certain operations. For this reason, in networks as complex as the one managed by Canal, the need arises to propose so-called "**topological observability**", which consists of analysing whether, for a given set of measurements, it is possible to know not only the hydraulic state of the network, but also the connections that occur between its elements and, in particular, the state of the pumps and valves in it.

Incorporating topological observability involves including the existence of the valves and pumps in the Jacobian matrix for the system, which will also need to be full rank for the system to be observable. Given that in the last analysis it is desired to know the state of each of these elements, it is necessary to introduce as many additional state variables are there are valves and pumps. Specifically, selected as additional state variables are the flow rates circulating through the valves ($Q_{V_{ij}}; \forall ij \in L^V$) and pumps ($Q_{B_{ij}}; \forall ij \in L^B$), where L^V and L^B are the set of valves and pumps present in the network, respectively. As such, the inclusion of the valves and pumps requires as many additional columns to be added as there are elements of this type existing.

On the other hand, these two types of elements are introduced to the system as pipes or line-type elements of zero length, with it being possible to measure the piezometric levels at their start and end nodes (which will be measurements that are available or candidates depending on the network measurement configuration). Furthermore, depending on whether or not information is available on the characteristic curve for the pump or valve, two approaches can be adopted:

- If **information on the characteristic curve is not available** for the particular element, it is not possible to establish any relationship between the flow rate circulating through the pump or valve and the piezometric levels at the end nodes. Additionally, as the flow rate circulating through a pump or valve cannot be measured in that element as such (although it could be measured in the pipe immediately before or following), rows corresponding to the flow rate in these particular elements are not included. Therefore, these two groups of additional columns have null values in all the flow rate measurements. In addition, they have zeros in the rows corresponding to piezometric level measurements, as it is assumed that no specific information is available on the characteristic curves or losses in the element. In turn, these additional state variables, which in the last analysis are flow rates, do intervene in the balances at the nodes, with their partial derivatives adopting values of:

  1 (output flow from the node) or

  -1 (input flow to the node) in this last block according to the expression **(13)**.

$$\boldsymbol{J}_t = \begin{bmatrix} \frac{\partial \tilde{h}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_n} & \frac{\partial \tilde{h}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{h}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{h}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{h}_1}{\partial Q_{B_{n_b}}} \\ \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{h}_n}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_n} & \frac{\partial \tilde{h}_n}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{h}_n}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{h}_n}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{h}_n}{\partial Q_{B_{n_b}}} \\ \frac{\partial \tilde{Q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_n} & \frac{\partial \tilde{Q}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_1}{\partial Q_{B_{n_b}}} \\ \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{Q}_{n_p}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_n} & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{B_{n_b}}} \\ \frac{\partial \tilde{q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_n} & \frac{\partial \tilde{q}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{q}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{q}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{q}_1}{\partial Q_{B_{n_b}}} \\ \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{q}_{n_q}}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial h_n} & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{B_{n_b}}} \end{bmatrix} \quad (14)$$

VALVES PUMPS

- If **information is available on the characteristic curve** for the particular element, it is possible to include an additional row relating to the flow circulating that establishes the relationship between the piezometric levels at both ends of the particular element (see expression **(15)**). Typically, this relationship is determined by a quadratic or parabolic equation, either to include the energy input to the system by the pump or the losses that occur through a certain valve.

This variety of possibilities means it is necessary to incorporate rows or columns depending on the information available and the type of each particular pump and valve, which demands versatility and consistency in the process of constructing the Jacobian matrix.

$$
\boldsymbol{J}_t = \left[\begin{array}{ccccc|ccc|ccc}
\frac{\partial \tilde{h}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_n} & \frac{\partial \tilde{h}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{h}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{h}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{h}_1}{\partial Q_{B_{n_b}}} \\
\vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
\frac{\partial \tilde{h}_n}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_n}{\partial h_n} & \frac{\partial \tilde{h}_n}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{h}_n}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{h}_n}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{h}_n}{\partial Q_{B_{n_b}}} \\
\frac{\partial \tilde{Q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_n} & \frac{\partial \tilde{Q}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_1}{\partial Q_{B_{n_b}}} \\
\vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
\frac{\partial \tilde{Q}_{n_p}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial h_n} & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{n_p}}{\partial Q_{B_{n_b}}} \\
\hline
\frac{\partial \tilde{Q}_{V_1}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{V_1}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{V_1}}{\partial h_n} & \frac{\partial \tilde{Q}_{V_1}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{V_1}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{V_1}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{V_1}}{\partial Q_{B_{n_b}}} \\
\vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
\frac{\partial \tilde{Q}_{V_{n_v}}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial h_n} & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{V_{n_v}}}{\partial Q_{B_{n_b}}} \\
\frac{\partial \tilde{Q}_{B_1}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{B_1}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{B_1}}{\partial h_n} & \frac{\partial \tilde{Q}_{B_1}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{B_1}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{B_1}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{B_1}}{\partial Q_{B_{n_b}}} \\
\vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
\frac{\partial \tilde{Q}_{B_{n_b}}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial h_n} & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{Q}_{B_{n_b}}}{\partial Q_{B_{n_b}}} \\
\hline
\frac{\partial \tilde{q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_n} & \frac{\partial \tilde{q}_1}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{q}_1}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{q}_1}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{q}_1}{\partial Q_{B_{n_b}}} \\
\vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
\frac{\partial \tilde{q}_{n_q}}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial h_n} & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{V_1}} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{V_{n_v}}} & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{B_1}} & \cdots & \frac{\partial \tilde{q}_{n_q}}{\partial Q_{B_{n_b}}}
\end{array}\right] \tag{15}
$$

VALVES PUMPS

5.3.3. Inclusion of sectors

To reproduce the reality of the network managed by Canal de Isabel II illustrated in Figure 1, it was proposed, additionally, to include the sectors and the dividing valves or connections between sectors in the hydraulic model and the subsequent observability analysis. As stated, the philosophy of sectorisation consists basically of having a main transport network (in this case the strategic network) that supplies water to independent sectors or DMAs, which have limited, controlled input and output points, for improved knowledge of flows in the network.

The incorporation of these elements makes it possible to reproduce the complete behaviour of the network subject to study and is motivated by the fact that the input flow rate measured in each sector does not necessarily correspond to actual network consumption, which can on the other hand be estimated (pseudo-measurement) by means of CHYPRE at practically any time instant. For this reason, it is of interest to identify the observable variables and the uncertainty associated with the flows in the connections existing between sectors as, if these flows were observable, it would be possible to know the inputs to and outputs from the various sectors in the subsequent state estimation process.

INDEX

The inclusion of sectors and connections between sectors in the hydraulic model and the subsequent analysis makes it necessary to take the particularities of flow that takes place in these zones into account. The sectors are in reality an artificial way of representing a whole zone of the system, hence these points represent aggregate demand values to which piezometric levels cannot be attributed. Therefore, only continuity or mass conservation equations operate at the nodes considered as sectors. In this way, it is necessary to separate the operation of the strategic network, which is being governed by the balance equations at nodes and energy loss in pipes, and the part corresponding to the sectors, where only continuity equations apply. This makes it possible to integrate the information into the Jacobian matrix, where as many additional state variables are defined as there are connections between sectors, although the measurements corresponding to the pseudo-measurements for consumption in the sectors and the knowledge of the flow between them partly offset the additional information necessary to characterise the flow in this zone.

6. Identification of observable islands in the strategic network

INDEX

This chapter firstly presents the mathematical approach for the methodologies adapted to identify observable islands in water supply systems, then evaluating these, to finally select the most suitable methodology with a view to implementing it in the strategic network managed by Canal. Once the most suitable method has been selected, the observable islands in the network under normal operating conditions are identified. Finally, specific analysis are performed to make it possible to identify the pumps, valves (including diversion pumps) and breaks that are observable under these conditions.

6.1. ADAPTATION OF THE METHODOLOGY TO THE STRATEGIC NETWORK

In the electricity sector, which is a point of reference in matters of observability and state estimation in networks, highly diverse methodologies have been developed for identifying observable islands. However, due to the limited studies of observability in water networks, few methodologies have been adapted to supply systems. Presented below are the methodologies considered at prototype scale in this research for possible use in identifying observable variables for the strategic network.

6.1.1. Adapted methodologies

The methodologies adapted and implemented for identifying observable islands in water supply systems are aimed at analysing the rank of the measurement Jacobian matrix for the system, which must allow identification of the observable variables for a given measurement configuration. The mathematical basis for each of the five methods considered is summarised below.

1) Null space method

In general, the null space N_A for a matrix A of dimensions $m \times n$ is denoted as:

$$N_A = \{x \in \Re^n : \ Ax = 0\} \qquad (16)$$

In this way, the dimension of the null space makes it possible to determine whether the matrix forms a determinate compatible system (the null space is an empty set) or whether, instead, it is of a certain dimension, indicating that several rows of the matrix are linearly dependent.

As the necessary condition for a water distribution to be observable is that the Jacobian matrix associated with the subset of measurements available is of full rank, this technique makes it possible to evaluate observability easily. To do this, it is necessary to extract the Jacobian matrix associated with the set of measurements available $\boldsymbol{J_a}$, given by the following expression:

INDEX

$$\boldsymbol{J_a} = \begin{bmatrix} \frac{\partial \tilde{h}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{h}_j}{\partial h_1} & \cdots & \frac{\partial \tilde{h}_j}{\partial h_i} & \cdots & \frac{\partial \tilde{h}_j}{\partial h_n} \\ \frac{\partial \tilde{Q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{Q}_{ij}}{\partial h_1} & \cdots & \frac{\partial \tilde{Q}_{ij}}{\partial h_i} & \cdots & \frac{\partial \tilde{Q}_{ij}}{\partial h_n} \\ \frac{\partial \tilde{q}_1}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_1}{\partial h_n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ \frac{\partial \tilde{q}_j}{\partial h_1} & \cdots & \frac{\partial \tilde{q}_j}{\partial h_i} & \cdots & \frac{\partial \tilde{q}_j}{\partial h_n} \end{bmatrix} \begin{matrix} ;\ \forall j \in \Omega_h \\ \\ ;\ \forall ij \in \Omega_Q \\ \\ ;\ \forall j \in \Omega_q \end{matrix} \tag{17}$$

Where:

- Ω_h Is the set of measurements of piezometric height.
- Ω_Q Is the set of flow measurements .
- Ω_q Is the set of measurements of demand or consumptions (including transit nodes).

In this way, it suffices to calculate the null space associated with $\boldsymbol{J_a}$ to know whether the system is observable (null space is empty set) or not (null space of certain dimension). Furthermore, the null space N_A provides information on which state variables are observable and which are not for a given non-observable scenario (Castillo et al., 2005). This makes it possible on occasions to detect the piezometric levels that are observable, but not the regions formed by observable variables in general, as it does not make it possible to evaluate the observability of the demands or the circulating flows (variables that are not state variables).

2) Optimisation method

This approach, presented for electricity networks by Habiballah and Irving (2001), is based on establishing a series of conditions corresponding to measurements of the piezometric levels, flow rates (loss equations) and demands (continuity equations) available in the system. This set of expressions are constraints for the optimisation problem and reproduce the equations governing the system and, therefore, the structure of the measurement Jacobian matrix available ($\boldsymbol{J_a}$). Additionally, it adds two additional constraints that make it possible to solve the system of equations for the network by minimising the sum of two additional variables p_{ij} and $h_{i,aux}$, which are introduced to the system such that the solution values remain as limited as possible.

The expression **(18)** covers the problem of linear programming to be solved to analyse the observability of a water distribution network:

$$
\begin{aligned}
&\min \sum_{ij=1}^{n_p} p_{ij} + \sum_{j=1}^{n} h_{i,aux} \\
&\text{with the following restrictions:} \\
&h_i - 1 = 0 && ;\forall i \in \Omega_h \\
&h_i - h_j = 0 && ;\forall ij \in \Omega_Q \\
&card_i \cdot h_i - \sum_{j=1}^{card_i} h_j = 0 && ;\forall i \in \Omega_q \\
&h_i - h_j + p_{ij} \cdot k = k && ;\forall ij \in L \\
&p_{ij} \geq 0 && ;\forall ij \in L \\
&h_{i,aux} \geq 0 && ;\forall i \in \Omega_h \\
&h_{i,aux} \geq h_i && ;\forall i \in \Omega_h
\end{aligned}
\qquad \textbf{(18)}
$$

Where $card_i$ refers to the number of pipes entering and leaving the node i, and k is a vector that covers the number of pipes existing in the system ($k = 1,\ 2, \ldots, n_p$).

Solving this linear programming problem gives the value adopted at the optimum. As can be deduced from the formulation given by the expression **(18)**, these variables take the value **1** when they are observable and a value **other than 1** when this is not the case. In this way, as in the case of the previous method, this proposal based on optimisation makes it possible to identify the state variables that are observable, but not the observable islands as such (it does not permit evaluation of the observability of flows and consumptions).

3) Algebraic method

This method was presented by Pruneda et al. (2010) for electricity networks and adapted by Díaz et al. (2015) to water supply networks. In outline, this methodology sets out the application of an elimination technique based on the Gauss method to the reordered Jacobian matrix for the network, containing firstly the available system measurements ($\boldsymbol{J_a}$, of dimension $m_a \times n$) and secondly the rest of the variables that, although they are not available, are candidate measurements that could be available at a certain cost ($\boldsymbol{J_c}$, of dimension $m_c \times n$). In this way, the resulting transformed matrix at the end of the algorithm, the detailed explanation of which is available in Díaz et al. (2015), contains all the information on system observability, and keeps the relationships between all the network's hydraulic variables up to date. This makes it possible to identify redundant measurements (whether available or candidates) that would guarantee the same level of observability in case of fault in any of them. The transformed matrix also makes it possible to identify critical measurements, with a critical measurement understood as that which is essential for the observability of the system, as well as to detect candidate measurements that could replace them in the case of a sensor fault, a fault in the telemetry system, etc.

However, the most outstanding aspect of this proposal is that it makes it possible to evaluate the observability of all the network variables, hence identification of observable islands can be performed when the system as a whole is not observable as, unlike the previous methods, it is not limited exclusively to evaluating the observability of the state variables.

4) Binary method

This method is presented in detail in electricity networks by Solares et al. (2009) with the aim of minimising rounding errors or numerical instabilities that could arise by applying algebraic techniques such as those used in the previous method, which work with some tolerance to identify null elements. As such, the sole difference with respect to the algebraic method consists of modifying the starting reordered Jacobian matrix by assigning the value 1 to elements other than zero and 0 to the rest. Subsequently, this matrix is transformed by means of a binary operation rule that is equivalent to the elimination technique used in the previous method.

The implementation of the binary version has potential as it speeds up the matrix transformation process. However, the need to replace non-null elements with a unique value of 1 in the starting Jacobian matrix implies a significant loss of information. This means that when there are a large number of elements that are not null, the binary operation rules cannot make the complete transformation, with unknown values resulting that cloud the relationships between measurements and state variables. This situation is resolved by implementing the algebraic method in certain sub-matrices of the original Jacobian matrix, but the need for an algebraic complement significantly slows the methodology.

This can be particularly problematic when working with known consumption values (whether measurements or pseudo-measurements), as the rows of the Jacobian matrix corresponding to demand measurements are associated with a large amount of non-null elements as a result of establishing relationships with the piezometric levels for all the end nodes that condition the balance in a given node (see equation **(13)**). The presence of multiple non-null elements in the matrix favours the propagation of the aforementioned unknown values, making it necessary to apply the algebraic complement in a large number of situations and for matrices of a similar size to the original. The appearance of these unknown values means that although the method is intended, like the algebraic method, to evaluate the observability of all the hydraulic variables, on many occasions only the observability of the state variables can be analysed with certainty. Furthermore, this loss of information limits the possibility of identifying observable islands.

5) Stochastic method

This method is described in detail in Díaz et al. (2016 b). It presents a different approach to the previous methods, as it employs the equations governing the system (balance and loss equations) to relate the uncertainty associated with the different variables instead of with the variables themselves.

The basis of the method is that in order to be able to resolve the uncertainty for any flow network, it is necessary to complement the n_n balance equations at the nodes that are not tanks $(\forall i \notin V^D)$ and the n_p energy loss equations applying to each of the pipes ($\forall ij \notin L$) with $n_n + n_d$ additional conditions that make it possible to make the number of equations ($n_n + n_p$) the same as the number of variables or unknowns ($2 \cdot n_n + n_d + n_p$), resulting from considering the piezometric levels at the nodes ($h_i; \forall i \in V$), the consumptions at the nodes that are not tanks ($q_i; \forall i \in (V^Q \cup V^T)$) and the flows in the pipes ($Q_{ij}; \forall ij \in L$). This approach is analogous to that for conventional solving of network flow, where it is sufficient to know the consumptions at the nodes (n_n) and the levels in the tanks (n_d) to be able to solve for the flow or, to put it another way, for the network to be observable.

In this way, the uncertainties for the various network variables, which are shown with an apostrophe (') alongside the variable in question, can be obtained by imposing $n_n + n_d$ additional conditions. In this case, it is assumed that the distribution of uncertainties is known for the consumptions at the nodes that are not tanks $q_i' \in N(0,\sigma_i)$ and the piezometric levels in the tanks $h_{d_j}' \in N(0,\sigma_j)$, which are Gaussian with a mean of zero. On these assumptions, the uncertainty of the known variables can be propagated to the rest, giving a first estimate of the uncertainty of the state variables by summing the diagonal elements of the resulting variance-covariance matrix for the piezometric levels ($\sum \sigma^2{}_{h_0}$).

It is true that this initial value alone does not provide decisive information, but if it is compared with the value this same sum takes for a given measurement configuration, it can provide relevant information in terms of observability by making it possible to analyse how this improves on adding measurements to the system.

In this line, the method considers that if a variable is measured with an instrument, its value is exact or at least an order of magnitude lower than the uncertainty assumed for levels in tanks and the consumption at the nodes, meaning that the effects of deterministic observability can be considered negligible. This means that if a vector $\boldsymbol{X}$ of measurements and a vector $\boldsymbol{Y}$ of non-measurements (the rest) is assumed, the properties of a conditioned multivariate normal distribution can be used to obtain the variance-covariance matrix for all the system variables conditioned to the existence of these measurements $\boldsymbol{R}_{Y|X}$:

$$\boldsymbol{\mu}_{Y|X} = \boldsymbol{\mu}_Y + \boldsymbol{R}_{Y \cdot X} \cdot \boldsymbol{R}_{X \cdot X}{}^{-1} \cdot (\boldsymbol{X} - \boldsymbol{\mu}_X) \quad (19)$$

$$\boldsymbol{R}_{Y|X} = \boldsymbol{R}_Y - \boldsymbol{R}_{Y \cdot X} \cdot \boldsymbol{R}_X{}^{-1} \boldsymbol{R}_{X \cdot Y} \quad (20)$$

Note that the diagonal of the variance-covariance matrix provides the variance of all the system variables for the measurement configuration analysed, hence this approach also makes it possible to identify observable islands. In addition, the ratio of the conditional variance and the initial variance for each variable provides an idea of the level of observability of the variable, which makes it possible to add nuances beyond the observability/lack of observability for the variable in question.

In this line, on $\boldsymbol{R_{Y|X}}$ the sum of the variances for the piezometric levels ($\sum \sigma^2{}_{h_C}$) can be calculated again to compare it with the previously obtained baseline value ($\sum \sigma^2{}_{h_0}$). Hence defining the *Stochastic Observability Index referred to piezometric levels* (SOI_h [9]) on the ratio between the two sums:

$$SOI_h = (1 - \frac{\sum \sigma^2{}_{h_C}}{\sum \sigma^2{}_{h_0}}) \cdot 100 \qquad \textbf{(21)}$$

A measure of observability is obtained for a given measurement configuration, where this index has the value 100% when the system is observable.

Similarly, the *Stochastic Observability Index referred to flow rates* (SOI_Q) can be calculated as:

$$SOI_Q = (1 - \frac{\sum \sigma^2{}_{Q_C}}{\sum \sigma^2{}_{Q_0}}) \cdot 100 \qquad \textbf{(22)}$$

Furthermore, the observability of the system could be evaluated by weighting the uncertainties according to the flows circulating through each pipe. A greater weight could be assigned to main network flows in this way.

The *Stochastic Observability Weighted Index referred to flow rates* ($SOWI_Q$ [10]) would then be:

$$SOWI_Q = (1 - \frac{\sum Q \cdot \sigma^2{}_{Q_C}}{\sum Q \cdot \sigma^2{}_{Q_0}}) \cdot 100 \qquad \textbf{(23)}$$

It should be emphasised that measurement redundancy situations cannot be taken into account with this stochastic observability procedure, as it would not be possible to invert the covariance matrix $\boldsymbol{R_X}$ in these cases. This means the comparison of the index would be interesting for non-observable scenarios in which successive measurements are added to see the effect each measurement has on general network observability, but it would not make it possible to quantify the improvement achieved by adding additional measures in an observable scenario.

---

[9] SOI: Stochastic Observability Index

[10] SOWI: Stochastic Observability Weighted Index

6.1.2. Selection of the optimum methodology

Having presented the five methods implemented to identify observable variables, each of these was evaluated objectively to identify the most suitable technique for implementation in the network managed by Canal.

The level of adaptation of the different methodologies was considered in this analysis in accordance with the following criteria:

a. Type of output information

The five methodologies proposed for observability analysis can be classified into two groups depending on the type of output information they provide. On the one hand are the techniques that make it possible to evaluate system observability and identify the state variables that are inside and outside observable islands when the entire network is not successfully made an observable island. Included in this group are the null space, optimisation and binary methods. The binary method is included because, although it has the potential to evaluate the observability of all the variables, the appearance of unknown values limits the usefulness in practice of this method, being especially restricted when working with consumptions as available measurements, as is the case of the network managed by Canal. On the other hand are the techniques that make it possible to evaluate the observability of the system and identify the observable and unobservable variables when an observable island does not cover the entire network, whether state variables (piezometric levels), flows or demands. As such, the techniques included in this second group, which include the algebraic and stochastic methods, are the best option from the point of view of the usefulness of the output information.

b. Possibility of detecting observable islands

Detection of observable islands is only possible with those techniques that make it possible to evaluate the observability of all the network variables, as if only information on the state variables is provided, the output information is always isolated information. For this reason, and according to the above criterion, the algebraic and stochastic methods are the best option in this regard.

c. Type of information resulting for the purpose of analysing the classification of critical and redundant information

Knowing critical and redundant measurements from the observability point of view is of interest because it can determine how to react in the case of the possible loss of a measurement. In this regard, the null space and optimisation methods do not enable the critical and redundant measurements to be identified, while the algebraic method does, with the qualification that in the binary method the appearance of unknown values limits the extraction of information. The stochastic method has a significant limitation in this regard, as it does not permit working with redundant measurements. For this reason, from the point of view of extracting critical and redundant information, the algebraic method is the best option.

d. Potential for selecting and prioritising the siting of new metering devices

In this respect, all the methods are considered to behave in the same way, as the optimum location of devices is not proposed in this study in terms of improving network observability (extending observable islands of the system), but rather in terms of reducing observation uncertainty (see the point “Strategic location of additional metering devices”), which is a more realistic approach to the level of instrumentation and the telemetry system existing in the network managed by Canal.

e. Computational cost

Computational cost is a key aspect for implementing the observability analysis in the strategic network managed by Canal, as it conditions the viability of knowing the observable islands of the network under any circumstance for subsequent use for state estimation techniques.

The response of each method in terms of calculation time was evaluated by means of implementation for mesh networks of different sizes, meaning that the computational cost trend could be analysed for each method. These were networks with the same structure in all cases (two tanks, two demand nodes and the rest transit nodes) and of successive sizes of 6, 12, 50, 100, 500 and 1,000 nodes. The computational cost associated with each of these examples for a random, non-redundant measurement configuration is shown in Table 4.

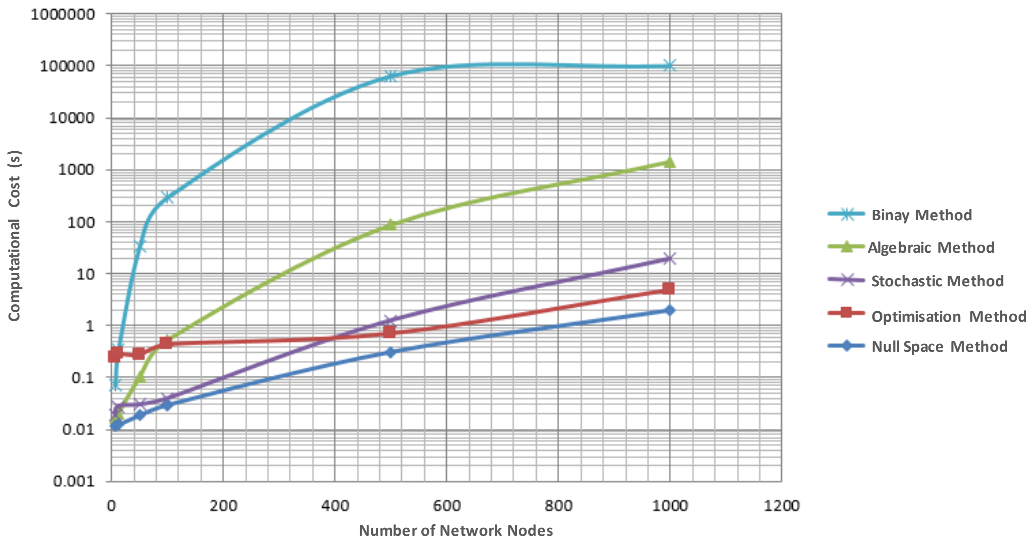
TABLE 4. COMPUTATIONAL COST OF IMPLEMENTATION IN MESH NETWORKS OF INCREASING SIZE

Activity	*Total computational cost (seconds)*					
	6 nodes	*12 nodes*	*50 nodes*	*100 nodes*	*500 nodes*	*1,000 nodes*
Flow network solution	0.2476	0.2476	0.2526	0.2739	9.5273	41,031
Generation of the Jacobian matrix	0.0077	0.0081	0.0132	0.0220	0.1793	0.6076
Null space method*	0.0040	0.0042	0.0054	0.0076	0.1277	1,3376
Optimisation method	0.2382	0.2839	0.2744	0.4355	0.7081	4.9521
Algebraic method*	0.0088	0.0124	0.0891	0.5063	86.861	1,449.7
Binary method*	0.0638	0.3402	32.705	292.26	63,628	> 100,000
Stochastic method	0.0187	0.0281	0.0301	0.0394	1.2504	19.319

** Methodologies that require the previous generation of the Jacobian matrix*

These results mean the binary method can be identified as being inefficient, as the need to perform a binary pivot to then implement the algebraic method does nothing more than slow down the process, increasing the computational cost considerably as the network size increases. However, the times shown in Table 4 also show the computational efficiency of the rest of methods, the cost of which remains more or less stable in spite of the progressive increase in network size. As shown in Figure 2, which reflects the variation in the computational cost of the method (including the generation of the Jacobian matrix when necessary) as a function of the number of network nodes, the adaptation of these methodologies as regards computational cost is, in descending order, the null space method, optimisation method, stochastic method and algebraic method, with the binary method discounted.

FIGURE 2. EVOLUTION OF COMPUTATIONAL COST WITH THE NUMBER OF NODES



f. Agility for combination with techniques for state estimation

The potential for applying observability techniques as a prior step complementary to state estimation is to a large extent determined by the agility of the method for considering each situation, which is highly dependent on the computational cost required to implement it. In this regard, and in view of the results above, the binary method has limitations, while the rest of the methods are suitable. It is interesting to mention the fact that the null space, optimisation and stochastic methods were highly efficient, with times of the order of seconds or tens of seconds in networks of 1,000 nodes. The algebraic method deserves special attention since, although it is less efficient, this proposal only requires some iteration to update the transformed matrix once this has been obtained initially and it is not necessary to confront the total computational cost each time. This fact makes the algebraic method particularly attractive from the point of view of implementation, as on storing all the information and relationships between the network variables, it makes it possible to progressively update the observable islands of the system to monitor their evolution.

In view of this analysis, the algebraic method shows itself to be the optimum methodology for identifying observable islands in the network managed by Canal, as not only is it efficient in terms of computational cost (which makes it viable for joint application with state estimation techniques), but it also makes it possible to evaluate the observability of all the variables (not just state variables) and it provides very full output information, making it possible to classify the measurements as redundant or critical for a given flow state in the network. For this reason, the algebraic method is selected for implementation in the strategic network in the following points.

Only one negative comment needs to be made with regard to this method, which is, as stated previously, that it requires the definition of a null or tolerance criterion, which must be defined in a consistent manner to avoid false identification of certain variables as observable when in reality they are not. This is an especially significant risk in large networks, where the accumulation of error can lead to discrepancies in the results. For this reason, implementing the method in the strategic network required a sensitivity analysis on this parameter (which in these examples in the prototype stage were considered as 10^{-6}), which was viable thanks to the possibility of comparing the results from the algebraic method with the rest of the low-computational-cost methods.

Finally, the potential of the stochastic method as a complement to the results from the algebraic method in scenarios that are not observable should be highlighted, as it enables the level of observability of each of the variables involved to be quantified beyond a “yes/no” response. However, in this work the option was taken to quantify observation uncertainty in any scenario, even if it is redundant, as the reduction of this uncertainty is the criterion selected for locating additional metering devices in chapter 8 “Strategic location of additional metering devices”.

6.2. IDENTIFICATION OF OBSERVABLE ISLANDS UNDER NORMAL OPERATING CONDITIONS

The algebraic method for the identification of observable islands makes it possible to evaluate whether sufficient algebraic relationships exist to calculate a given hydraulic variable based on the set of available measurements. To do this, this technique is first applied to the open zone of the network according to the network configuration or topology under normal operating conditions. This makes it possible to evaluate the observability/lack of observability of the network variables with the valves and pumps in the position associated with the habitual system operation, which is known by Canal. For this analysis, the *pseudo-measurements* for consumption in the sectors are considered to be known, as the information provided by the CHYPRE system provides redundancy to certain metering devices and is key to characterising the flow between sectors.

Under these assumptions, the strategic network as a whole is an **observable island**. This is explained because the demands that move the water in the network are known, with these being measured in the main at the outlets to sectors and as *pseudo-measurements* in those sectors that do not yet have flow meters at their inlet, and the network parameters linked to modelling hydraulic losses, the state of the valves and the characteristic curves for the system pumps are assumed to be known and, finally, the values calculated by the hydraulic model are accepted as correct. In this way, the observation uncertainty can be calculated in this scenario (see the chapter “Quantification of the observation uncertainty in the strategic network”) and the state of the entire network can be estimated.

However, it is unlikely that normal operating conditions are maintained at all times. Therefore, the identification of observable elements becomes important for characterising scenarios in which deviations occur from the assumed theoretical conditions. Specifically, in this research the choice was made to evaluate how system observability would be affected in the following situations:

1) That the operation of the system pumps and regulating valves change their operation or their assumed characteristic curve is no longer adequate.

2) That the diverting valves that are closed under normal conditions are opened, impeding flow to other network zones.

3) That breaks in the system are intended to be detected[11].

To tackle these three cases, an individual analysis of the "fault" or lack of knowledge of each of the elements involved in the three cases above was performed, which were each of the network pumps and valves, the demand for the closure nodes or neighbouring nodes located immediately before the closed diverting valves, and the transit nodes in the strategic network, respectively. In all these cases, individual changes were made to the elements in question and the state of the rest of the system was assumed to be according to the normal operating conditions.

It is important to emphasise that the analysis presented below were intended to evaluate the observability of the various network elements in case of individual variations in the topology occur. It is true that situations can occur in which there is more than one difference at the same time with respect to the expected theoretical scenario, but to weigh up all the possible combinations would be unmanageable and it is also impractical, as it will never be possible to measure everything. The simulation of the individual loss of certain elements does, however, give an idea of the robustness of the system in terms of observability. In this way, if in practice more than one difference does occur with respect to the theoretical scenario corresponding to the state of pumps and valves for normal operating conditions, it would be possible to detect at least the existence of a fault in the zone, although it may not be possible to determine in which element it occurred.

6.3. IDENTIFICATION OF OBSERVABLE PUMPS AND VALVES

In this scenario, the effect of assuming the state of the system valves and pumps to be known or unknown is evaluated. To do this, it was studied whether, on doubting the state of pumps and valves, these elements are in an observable island or not, by performing an analysis in which the state of each pump and valve are assumed to be unknown individually (one by one analysis), as in the habitual scenario they are all assumed to be known in accordance with habitual network operation. It is intended in this way to test whether, not knowing the state of a given pump or valve, its state can be observed based on the rest of the measurements available in the network. To do this, the state of the rest of the pumps and valves was assumed as per the normal operating conditions in the strategic network and the pseudo-measurements were considered to be known. The summary of the results is shown in Table 5.

[11] The connections between sectors are always algebraically observable as there are *pseudo-measurements*. However, it needs to be determined whether they are observable in practice, paying attention to the uncertainties.

TABLE 5. PUMPS AND VALVES OBSERVABLE ON DOUBTING THEIR OPERATION INDIVIDUALLY

Variables	*Total nº of elements*	*Total nº of observable elements*	*% of observable elements*
Shut-off valve flow	6,514	3,431	53
Regulating valve flow	507	351	69
Pump flow	123	100	81

According to these results, the state (open/closed), and consequently the flow through the shut-off valves in the open zone of the network may be observed at a percentage of around 50%, provided that their state is doubted individually. In the case of regulating valves, their flow is algebraically observable in 69% and in pumps this rises to 81%. These results show that in a significant proportion of the cases in which the state of a valve or pump is doubted, it would be possible to determine the state of the singular element. It should be said that this is a necessary condition, given that the element has sufficient relationships for its state to be determined. A second subsequent condition to be met, for specific operating conditions, is the uncertainty of the estimate of operation that can be made. However, this first condition is met in a very large number of cases, which shows the potential the measurement system provides for knowing the state of the network.

Note that evaluating this reality is possible thanks to the concept of topological observability presented in this work, which makes it possible to evaluate not only whether or not the variables are observable, but also whether it would be possible to determine the state of the pumps and valves based on the available measurements. As such, this analysis makes it possible to identify the pumps and valves in which it would be possible to detect unnotified operations in a subsequent state estimation process. It is true that these results limit themselves to individual variations with respect to the normal operating conditions, but bearing in mind that state estimation is a process intended to be implemented in real time, excessive alterations are not to be expected between one time step and the next. In any case, if simultaneous operations do occur, it may not be possible to correctly identify the valve or pump whose state is different, even though the elements were observable in this one by one analysis. However, it would be possible to detect that the network topology in the zone is not as expected.

6.4. IDENTIFICATION OF OBSERVABLE DIVERTING VALVES

This analysis is intended to evaluate the effect of changes in operation or operating the valves and pumps that regulate the diversions, i.e. those elements from which a closed zone of the network is detected, with flow unable to enter it due to a valve or pump being closed or off, respectively. To do this, an individual analysis was performed of demand at the so-called closure nodes or neighbouring nodes, which are those immediately before the pump or valve that closes the diversion.

As such, the observability/lack of observability of the closure nodes was evaluated by performing an analysis in which its demand was assumed to be unknown individually (one by one analysis). In this way, it can be evaluated whether the flow entering that region would be observable in case the valve or pump in question was opened or started up, i.e. the diversion was opened. To do this, the state of the rest of the pumps and valves was assumed as per the normal operating conditions in the strategic network and the pseudo-measurements were considered to be known.

The results of this analysis are presented in Table 6, where it can be seen that 42% of these neighbouring nodes would be observable. This means that in the event of a flow diversion through these nodes (always in an isolated manner, one by one), it would be possible to observe this from the algebraic point of view in 42% of cases. A different aspect is the precision with which it would be observed, which could mean that small flows are not in fact observable when the observation uncertainty is estimated.

TABLE 6. **DIVERTING VALVES OBSERVABLE ON DOUBTING THEIR OPENING INDIVIDUALLY**

Variables	*Total nº of elements*	*Total nº of observable elements*	*% of observable elements*
Closure node demand	2,380	1,003	42

6.5. IDENTIFICATION OF OBSERVABLE PIPE BURSTING

This scenario evaluates the observability/lack of observability of the network transit nodes that are not closure nodes (considered in the previous section) by performing an analysis in which it is doubted that the demand at these nodes is zero individually (one by one analysis). This makes it possible to evaluate whether it would be possible to detect that a non-zero demand is occurring at that node based on the available measurements, which would permit breaks to be detected. To do this, the state of the rest of the pumps and valves was assumed as per the normal operating conditions in the strategic network and the pseudo-measurements were considered to be known.

The summary of the observable elements identified in this case appears in Table 7, where it can be seen that only 9% of the transit nodes would be observable from the algebraic point of view in this case. Furthermore, it must be borne in mind that for those that are observable, the precision with which their value is estimated may mean that operationally a given break is not observable due to the associated uncertainty.

TABLE 7. **BREAKS OBSERVABLE ON DOUBTING THE DEMAND OF TRANSIT NODES INDIVIDUALLY**

Variables	*Total nº of elements*	*Total nº of observable elements*	*% of observable elements*
Transit node demand	13,702	1,210	9

These results show that the capacity to detect breaks is limited, as doubting the possibility of a diversion of unknown flow rate in the transit nodes, which constitute a significant part of the strategic network managed by Canal de Isabel II, represents a significant loss of information. To this end, a specific study should be conducted on the ability to detect breaks and/or leaks.

7. Quantification of the observation uncertainty in the strategic network

INDEX

Quantifying uncertainty is the second stage of the observability study conducted in this research. This phase makes it possible to quantify the uncertainty in the observation of the hydraulic variables existing in the network for a given scenario. Specifically, work was carried out with the flow scenarios considered in the hydraulic simulations which Canal has available for maximum, mean and minimum demands, also assuming knowledge of the pseudo-measurements to complement the measurements available in the network and so guarantee system observability.

The standard deviations considered for the metering devices depend on their precision, as shown in Table 3. In turn, real-time implementation of observability analysis would provide estimates of consumption or pseudo-measurements from CHYPRE, a computer application for estimating demands in sectors, available in Canal. However, to reproduce the high levels of uncertainty associated with estimates of this kind, a standard deviation of 20% of the mean value for the consumption in question is considered in this research, which means a significantly higher deviation than that of the metering devices themselves, such as the 2% or 5% assumed for flow measurement in pipes using flow meters.

Shown below is the detail of the methodology adapted to the strategic network to quantify the uncertainty, as well as a summary of the results obtained for the maximum, mean and minimum demand scenarios.

7.1. ADAPTATION OF THE METHODOLOGY TO THE STRATEGIC NETWORK

The methodology traditionally adopted for quantifying observation uncertainty is included under the point entitled "Quantification of observation uncertainty", based on estimating the uncertainty of state estimation understood as a least squares problem. However, this proposal involves solving a least squares problem in which weights are assigned to the various available measurements. For example, the measurement of consumption at transit nodes is associated with a deviation very close to zero, as it is known with certainty that the demand is zero, while the pseudo-measurements for consumption in the sectors are associated with large uncertainties. These order of magnitude differences in $\boldsymbol{R}_z$ can lead to the appearance of significant numerical errors in the solution of **(6)**, especially when working with a network of the dimensions of the strategic network managed by Canal, where these differences of magnitude in the weights are added to the differences in flow conditions to begin with.

For this reason, and with the idea of minimising the numerical errors associated with the measurement precisions, it was chosen to quantify the observation uncertainty by transforming the original unconstrained state estimation problem into a problem with constraints of the hydraulic type (see Díaz et al., 2016 a). In this way, those measurements without uncertainty, as they are known in the system (such as those associated with demand at a transit node, which is null), will be introduced to the problem as one constraint or equation more to be satisfied, so avoiding the need to assign a practically zero deviation that can lead to the appearance of the aforementioned numerical errors. Additionally, this approach makes it possible to introduce upper and lower limits to the variables, which is of interest, for example, for setting bounds for the piezometric levels in the tanks.

The statement of the state estimation understood as an optimisation problem with constraints is then:

$$\underset{x}{Min}\, F(\boldsymbol{x}) = \varepsilon^T \boldsymbol{R}_z^{-1} \varepsilon = [\boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x})]^T \boldsymbol{R}_z^{-1} [\boldsymbol{z} - \boldsymbol{g}(\boldsymbol{x})] \quad \textbf{(24)}$$

With:

$$\boldsymbol{f}(\boldsymbol{x}) = 0, \quad \textbf{(25)}$$

$$\boldsymbol{g}(\boldsymbol{x}) \leq 0. \quad \textbf{(26)}$$

Where **(25)** reproduces the hydraulic equality constraints or measurements not subject to error and **(26)** the inequality constraints representing the lower and upper limits for the variables. This statement of hydraulic state estimation could be solved using mathematical programming techniques, but the interest in this case lies in using this concept to quantify the uncertainty of state estimation and not in its solution in itself.

The quantification of uncertainty is a local analysis of the optimal solution $\hat{\boldsymbol{x}}$, hence once the optimum state estimation is known, the active inequality constraints are considered as equality constraints and the inactive are discarded. In this way, the vector $\boldsymbol{f}(\boldsymbol{x})$ contains p equality constraints and q_Λ inequality constraints, where Λ is the set of active inequality constraints.

Therefore, the first-order optimality conditions for the problem **(24)** - **(26)** at the optimum $\hat{\boldsymbol{x}}$ are:

$$\begin{gathered} \sum_{i=1}^{m} \nabla_{\boldsymbol{x}} [\omega_i (z_i - g_i(\hat{\boldsymbol{x}}))^2] + \sum_{i=1}^{p} \lambda_i \nabla_x f_i(\hat{\boldsymbol{x}}) = 0 \\ f_i(\hat{\boldsymbol{x}}) = 0, i = 1,\ldots, p + q_\Lambda \end{gathered} \quad \textbf{(27)}$$

Where $\boldsymbol{W} = \boldsymbol{R}_z^{-1}$ is a diagonal matrix of dimension $m \times m$ that contains the weights for the measurements ω_i, $\boldsymbol{F} = \nabla_x f(\hat{\boldsymbol{x}})$ is the equality constraint Jacobian matrix (of dimension $(p + q_\Lambda) \times n$) and $\boldsymbol{\lambda}_i$ is the Lagrange multiplier vector associate with the equality constraints derived from **(25)** - **(26)**, of dimension $(p + q_\Lambda) \times 1$.

Although it is assumed that the optimum solution is known throughout this process of quantifying the uncertainty of the state estimate (the result of solving the flow network can be assumed directly), the solution of the state estimation problem **(24)** - **(26)**, could also be performed by applying Newton's method to **(27)**, solving the following series of linear equations iteratively:

$$\begin{bmatrix} \boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{J} & \boldsymbol{F}^T \\ \boldsymbol{F} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Delta \hat{x}}_{(\nu+1)} \\ -\boldsymbol{\lambda}_{(\nu+1)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{\Delta \hat{z}}_{(\nu)} \\ -\boldsymbol{f}(\hat{\boldsymbol{x}}_{(\nu)}) \end{bmatrix} \quad \textbf{(28)}$$

Where $\boldsymbol{\Delta \hat{z}}_{(\nu)} = \boldsymbol{z} - \boldsymbol{g}(\hat{\boldsymbol{x}}_{(\nu)})$, that makes it possible to obtain an optimum such that $\boldsymbol{\Delta \hat{x}}_{(\nu+1)} = \boldsymbol{0}$, and $\boldsymbol{f}(\hat{\boldsymbol{x}}_{(\nu)}) = \boldsymbol{0}$. Note that the system coefficient matrix is denominated $\boldsymbol{U}$ hereinafter.

As regards quantifying uncertainty, it should be stressed that when equality or inequality constraints exist, these form a part of $\boldsymbol{F}$, hence it is not certain that $[\boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{J}]$ is an invertible matrix even though the system is observable.

For this reason, this constrained statement requires the calculation of the inverse of the matrix $\boldsymbol{U}$ to then select the part that establishes a linear relationship between $\boldsymbol{\Delta\hat{x}}_{(\nu+1)}$ and $\boldsymbol{\Delta\hat{z}}_{(\nu)}$, in this case the upper left quadrant $\boldsymbol{E}_1$:

$$\begin{bmatrix} \boldsymbol{\Delta\hat{x}}_{(\nu+1)} \\ -\boldsymbol{\lambda}_{(\nu+1)} \end{bmatrix} = \boldsymbol{U}^{-1} \begin{bmatrix} \boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{\Delta\hat{z}}_{(\nu)} \\ -\boldsymbol{f}(\boldsymbol{\hat{x}}_{(\nu)}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{E}_1 & \boldsymbol{E}_2^T \\ \boldsymbol{E}_2 & \boldsymbol{E}_3 \end{bmatrix} \begin{bmatrix} \boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{\Delta\hat{z}}_{(\nu)} \\ -\boldsymbol{f}(\boldsymbol{\hat{x}}_{(\nu)}) \end{bmatrix} \quad (29)$$

In this way, the alternative equation to **(5)** with this approach is:

$$\boldsymbol{\Delta\hat{x}} = \boldsymbol{E}_1 \boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{\Delta\hat{z}} = \boldsymbol{S}_{xz} \boldsymbol{\Delta\hat{z}}\,, \quad (30)$$

Where $\boldsymbol{S}_{xz}$ represents the state variable sensitivity matrix $\boldsymbol{x}$ with respect to the measurements $\boldsymbol{z}$. As such, the expression that makes it possible to calculate the variance-covariance matrix of the state variables $\boldsymbol{R}_{\hat{x}}$ in this case is:

$$\boldsymbol{R}_{\hat{x}} = \boldsymbol{S}_{xz} \boldsymbol{R}_z \boldsymbol{S}_{xz}^T\,. \quad (31)$$

Once the variance-covariance matrix of the state variables is known, the uncertainty of the rest of the variables can be obtained by means of the Equation **(7)**.

It is important to emphasise that although the expressions above provide exact formulas for calculating uncertainty, their application to large networks which have elements of very different magnitudes means that the matrix $\boldsymbol{U}$ to be inverted is ill conditioned on occasions. This difficulty was overcome in this work by applying Singular Value Decomposition (SVD) techniques, which make it possible to detect which of the calculated uncertainties are numerical noise-free and therefore are reliable.

The percentage validity of the calculation of uncertainty in the case of the network managed by Canal is high (see Table 12), as it is always over 75%, with percentages of validity of over 95% recorded in a significant proportion of the variables studied.

7.2. UNCERTAINTY FOR MAXIMUM DEMANDS

The summary of the results for this scenario of maximum demands is presented in Table 8, Shown in it are the total number of elements subjected to analysis for the different variables and the associated standard deviations σ. Note that for the piezometric levels the uncertainty (σp) is evaluated in metres, while for the flows and demands in the strategic network (σQ) a percentage with respect to the measured value is used, i.e. In coefficient of variation terms.

In order to give an idea of the importance of the uncertainty of the flows in the sector connections (also σ_Q), the uncertainty of the flow rate in the sector connections with the strategic network and between the sectors themselves is presented in litres per second.

TABLE 8. ESTIMATION UNCERTAINTY IN THE MAXIMUM DEMANDS SCENARIO

Variables	***Total no. of elements***				
Tank piezometric level	272				
Variables	***Total no. of elements***	***$\sigma_p \leq 1$ m***		***1 m < $\sigma_p \leq 10$ m***	***$\sigma_p > 10$ m***
Node piezometric level	32,493	30,298		1,463	732
Variables	***Total no. of elements***	***$\sigma_Q \leq 2\%$***	***2% < $\sigma_Q \leq 5\%$***	***5% < $\sigma_Q \leq 10\%$***	***$\sigma_Q > 10\%$***
Node demand	720	368	251	29	72
Pipe flow	26,015	19,525	3,583	1,063	1,844
Shut-off valve state	7,871	5,877	1,080	334	580
Regulating valve state	300	250	20	12	18
Pump state	159	103	9	9	38
Sector demand	669	173	149	50	297
Variables	***Total no. of elements***	***$\sigma_Q \leq 1$ l/s***	***1 l/s < $\sigma_Q \leq 5$ l/s***	***5 l/s < $\sigma_Q \leq 10$ l/s***	***$\sigma_Q > 10$ l/s***
Flow between sectors and strategic network	772	500	246	25	1
Flow between sectors	132	17	44	30	41

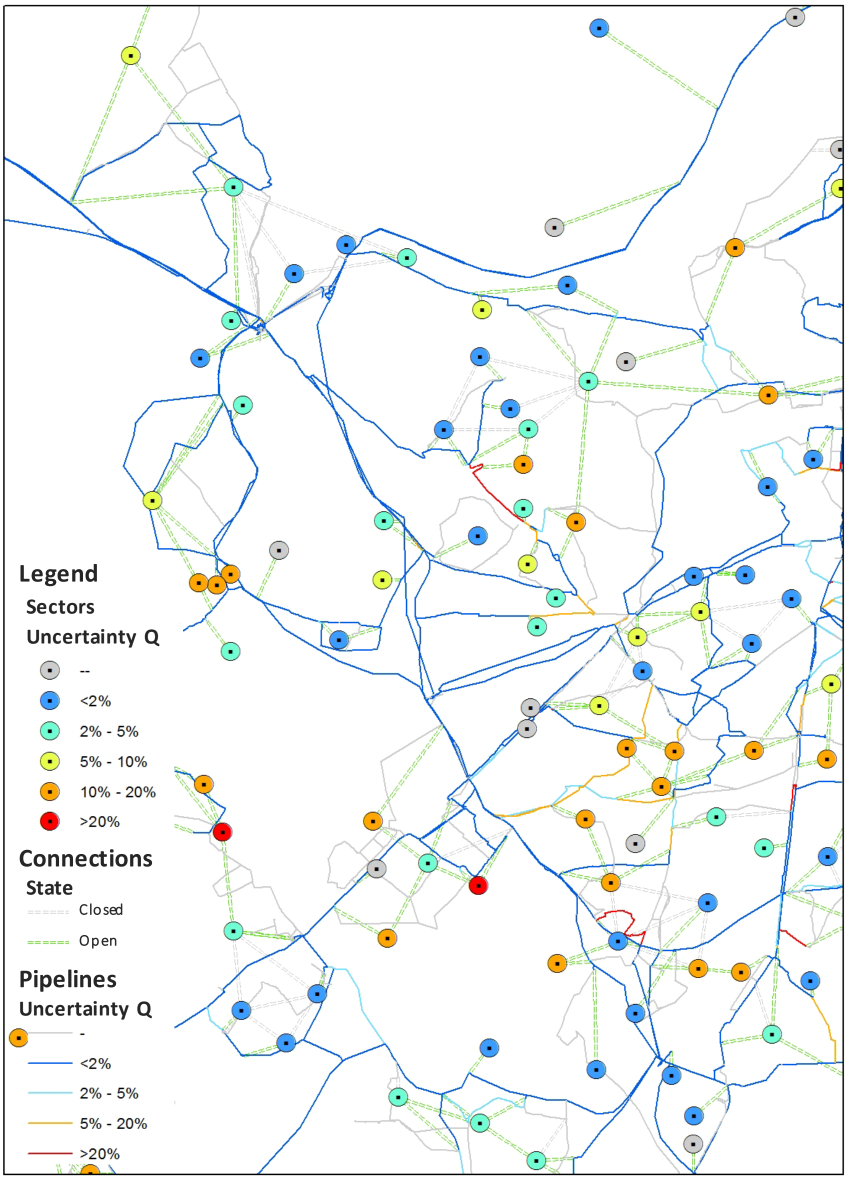
As can be seen in the results, a high percentage of the pressures can be monitored with an acceptable precision, although it must be borne in mind that this result is subject to considering the equations governing the hydraulic problem and, especially, the parameters that define it to be correct, which are not questioned in this work. In terms of flow rates in network elements, a significant majority are below one standard deviation in the measurement error of 5% of the expected value, which is reasonable according to the precision of the existing flow meters and the pseudo-measurements assumed for the consumptions in the sectors.

Finally, in the case of the connections, those between sectors are of interest. The precision of the estimates is very disparate in these, limiting the ability to identify on the basis of which flow rate it would be possible to detect the opening of the dividing valves separating the sectors. This uncertainty could make it possible to detect the opening of a sector provided that a flow rate occurs at that opening of, at least, twice the standard deviation of the estimate (with a significance level of 5% or possibility of a false positive). As such, considering the current measurement system could detect the opening of a sector depends to a great extent on the flow rates that can physically be diverted through each connection.

INDEX

In addition to the summary of results presented in Table 8, the implementation of the methodology for quantifying uncertainty makes it possible to characterise the precision with which the observable variables would be observed in the strategic network in this scenario. As proof of this, Figure 3 shows an example of the results obtained in terms of uncertainty in the sectorised strategic network for this scenario of maximum demand. It can be seen in them how the uncertainty for the demands in the sectors is considerably higher than that for the rest of the network, as the flows through the connections are determined, in the last analysis, by the estimates of consumption given by CHYPRE.

FIGURE 3. **RESULTS OF THE EVALUATION OF UNCERTAINITY IN THE SECTORISED DISTRIBUTION NETWORK**



Example of results of evaluating uncertainty in the sectorised distribution network, with a colour scale, where the circular symbols represent sectors and the dotted elements inlets to strategic network sectors from pipelines, which are represented as continuous elements

INDEX

7.3. UNCERTAINTY FOR MEAN DEMANDS

The results for this mean demand scenario are summarised in Table 9, where a very similar behaviour to the previous case can be seen, with the associated relative errors increased as the circulating flow rates are lower as a result of the reduction in demands.

TABLE 9. ESTIMATION UNCERTAINTY IN THE MEAN DEMANDS SCENARIO

Variables	*Total no. of elements*				
Tank piezometric level	270				
Variables	***Total no. of elements***	***$\sigma_p \leq 1$ m***		***1 m < $\sigma_p \leq 10$ m***	***$\sigma_p > 10$ m***
Node piezometric level	31,354	27,953		2,150	1,251
Variables	***Total no. of elements***	***$\sigma_Q \leq 2\%$***	***2% < $\sigma_Q \leq 5\%$***	***5% < $\sigma_Q \leq 10\%$***	***$\sigma_Q > 10\%$***
Node demand	724	255	302	64	103
Pipe flow	25,668	16,659	4,473	1,925	2,611
Shut-off valve state	7,779	4,954	1,357	619	849
Regulating valve state	281	224	25	11	21
Pump state	159	68	24	14	53
Sector demand	653	157	105	60	331
Variables	***Total no. of elements***	***$\sigma_Q \leq 1$ l/s***	***1 l/s < $\sigma_Q \leq 5$ l/s***	***5 l/s < $\sigma_Q \leq 10$ l/s***	***$\sigma_Q > 10$ l/s***
Flow between sectors and strategic network	769	600	163	6	0
Flow between sectors	132	44	41	22	25

7.4. UNCERTAINTY FOR MINIMUM DEMANDS

The results for this scenario of minimum demands are summarised in Table 10, In general, the same trend as in the previous case is maintained, with the relative uncertainties increased, although the absolute uncertainties decrease. High levels of precision (deviations < 5%) continue to be maintained for around 70% of the network elements. The same is not true of the sectors, which are measured with a precision of worse than 10% in around 50% of the elements. This indicates that, although the measurements exist, they do not provide sufficient precision as occurs in the case of the transport sectors.

TABLE 10. **ESTIMATION UNCERTAINTY IN THE MINIMUM DEMANDS SCENARIO**

Variables		***Total no. of elements***			
Tank piezometric level		269			
Variables	***Total no. of elements***	$\sigma_p \leq 1$ *m*		*1 m <* $\sigma_p \leq 10$ *m*	$\sigma_p > 10$ *m*
Node piezometric level	33,230	24,083		3,697	5,450
Variables	***Total no. of elements***	$\sigma_Q \leq 2\%$	$2\% < \sigma_Q \leq 5\%$	$5\% < \sigma_Q \leq 10\%$	$\sigma_Q > 10\%$
Node demand	724	253	229	72	170
Pipe flow	25,672	12,956	5,396	2,332	4,988
Shut-off valve state	7,785	3,828	1,620	762	1,575
Regulating valve state	281	224	25	11	21
Pump state	274	180	41	17	36
Sector demand	648	162	93	51	342
Variables	***Total no. of elements***	$\sigma_Q \leq 1$ *l/s*	*1 l/s <* $\sigma_Q \leq 5$ *l/s*	*5 l/s <* $\sigma_Q \leq 10$ *l/s*	$\sigma_Q > 10$ *l/s*
Flow between sectors and strategic network	764	685	76	3	0
Flow between sectors	129	75	43	9	2

INDEX

8. Strategic location of additional metering devices

INDEX

This chapter covers the identification of the most appropriate locations for installing additional metering devices with a view to improving the result of a subsequent state estimation. As stated under the point “Strategic location of additional metering devices”, in this work the option was taken of detecting the most interesting positions with a view to reducing the uncertainty of state estimation, as although the network has a good level of instrumentation, which can make it possible to observe quite a lot of topological conditions in it, the uncertainty obtained for the demand levels considered can be improved.

As such, this chapter presents the methodology adopted to propose the most advisable locations given a certain level of investment and the positions obtained for the specific case of the strategic network managed by Canal, as well as an analysis of these results.

8.1. ADAPTATION OF THE METHODOLOGY TO THE STRATEGIC NETWORK

The methodology for selecting the most suitable locations for installing additional metering devices was adapted to the network managed by Canal, dividing this into two large blocks:

1. Shortlisting locations for possible additional devices

This was performed based on the results obtained in the analysis of uncertainty for the average scenario (“Uncertainty for mean demands” point). Specifically, considered as possible locations were the pipes (pipes other than pumps, valves and connections between sectors) associated with the highest coefficients of variation among those that have higher speeds, associated with the most significant flows. In addition, these pipes needed to have a minimum length to guarantee correct device measurement.

Once the flows in the pipes capable of being measured were selected, according to the above criteria, the possible locations for devices were extended to other network nodes and pipes with a criterion of closeness and distance. As such additional devices may be installed in pipes and nodes neighbouring those selected that were not too far away. This strategy means that it is not necessarily those initially shortlisted locations that are selected for adding devices, but they can be located nearby, so that the methodology makes the best use of the relationships between the variables.

Note this the starting point is that all the outlets to the sectors (demands from the strategic network) have flow meters, as although they were not all installed when the study was performed, this work was already planned and is being carried out by Canal with the aim of completing network sectorisation.

2. Identification of the best locations

Once the candidate measurements have been shortlisted, the calculation of the inverse of the matrix $\boldsymbol{U}$ (according to equation **(29)**) can be dispensed with, implicitly or indirectly setting out in its place a reduction in uncertainty by taking advantage of the fact that the sum of the diagonal of the variance-covariance matrix of the state variables is equal to the eigenvalues of the matrix $\boldsymbol{U}$ and is therefore a measure of overall system uncertainty. Therefore, instead of minimising the eigenvalues of $\boldsymbol{U}^{-1}$ by summing them, it is proposed to maximise the sum of the inverse of the eigenvalues by means of the trace of $\boldsymbol{J}^T \boldsymbol{R}_z^{-1} \boldsymbol{J}$. This strategy makes it possible to reduce overall system uncertainty indirectly and set out the resulting problem as mixed-integer-linear, speeding up its solution.

8.2. IMPROVING THE OBSERVABILITY UNCERTAINTY FOR DIFFERENT INVESTMENT SCENARIOS

As argued previously, the consideration of pseudo-measurements in sector consumptions and tank levels leads to the entire network being in an observable island. This provides an initial level of observability, quantified in terms of the uncertainty linked to each of the variables in the problem. The inclusion of metering devices, up to the number currently placed, represents an increase in redundancy and an improvement in the precision of the state estimation that can be produced. On adding a larger number of devices, the uncertainty can be progressively, asymptotically reduced, up to the ideal limit of achieving a zero uncertainty.

This section contains the results of applying the strategy described above for the location of additional metering devices and evaluates how different investments influence the improvement in observation precision. To do this, four independent investment steps were considered as variants on the current situation. As such, this section will give the most advisable locations for installing additional devices at each level of investment, taking the associated direct costs (average price of the devices to be made available) and the installation work required (civil works, installation and communications) into account.

To analyse the effect of placing devices at the proposed locations, a series of aggregate parameters were used for both the current situation (uncertainty for the mean flow rate scenario) and for updated measurement configurations for each investment step. Specifically, the results obtained for each level of investment were compared using the following parameters representative of the strategic network, the majority of which are presented in coefficient of variation terms, which is the ratio of the standard deviation of the variable with respect to the mean value:

INDEX

- Mean values of the coefficients of variation associated with demands.
- Mean values of the uncertainty (standard deviation) associated with the piezometric heights of the nodes.
- Mean values of the coefficients of variation associated with the flows in pipes.
- Mean values of the coefficients of variation associated with the flows in pumps.
- Mean values of the coefficients of variation associated with the flows in valves.
- Mean values of the coefficients of variation associated with demands in the sectors.
- Mean values of the uncertainty (standard deviation) associated with the flows between sectors and between the sectors and the strategic network.

The values of these parameters are shown in Table 11 for each of the scenarios considered. Note that this table also includes the value of the coefficient of variation and/or the measurement uncertainty in the so-called baseline case, corresponding to the theoretical situation in which only pseudo-measurements are available, which are associated with an uncertainty of 20% in coefficient of variation in the demands (both in the strategic network and in the sectors it supplies) and a deviation in the piezometric levels in the tanks of 1 m. This baseline or reference case means the worst case uncertainty scenario can be established, which is the opposite extreme to the ideal case, in which the coefficients of variation and/or uncertainties for all the variables would be zero as a result of the existence of error-free measurements.

TABLE 11. **AGGREGATE RESULTS OF UNCERTAINTY WITH DIFFERENT LEVELS OF INVESTMENT**

Aggregate results	***Level of investment***					
	Baseline *(pseudo-measurements)*	***Current***	***1***	***2***	***3***	***4***
Coefficient of Variation for demand (%)	17.23	10.18	9.00	8.92	8.77	8.60
Uncertainty in piezometric height (m)	6.30	1.30	1.14	1.13	1.13	1.08
Coefficient of Variation for flow in pipes (%)	11.93	4.17	3.73	3.56	3.52	3.49
Coefficient of Variation for flow in pumps (%)	10.31	3.32	3.04	3.04	3.04	2.92
Coefficient of Variation for flow in valves (%)	18.79	1.92	1.72	1.67	1.64	1.66
Coefficient of Variation for demand in sectors (%)	10.89	8.18	6.04	6.34	6.31	6.34
Uncertainty in flow between sectors (l/s)	3.21	2.28	1.02	1.16	1.16	1.16

INDEX

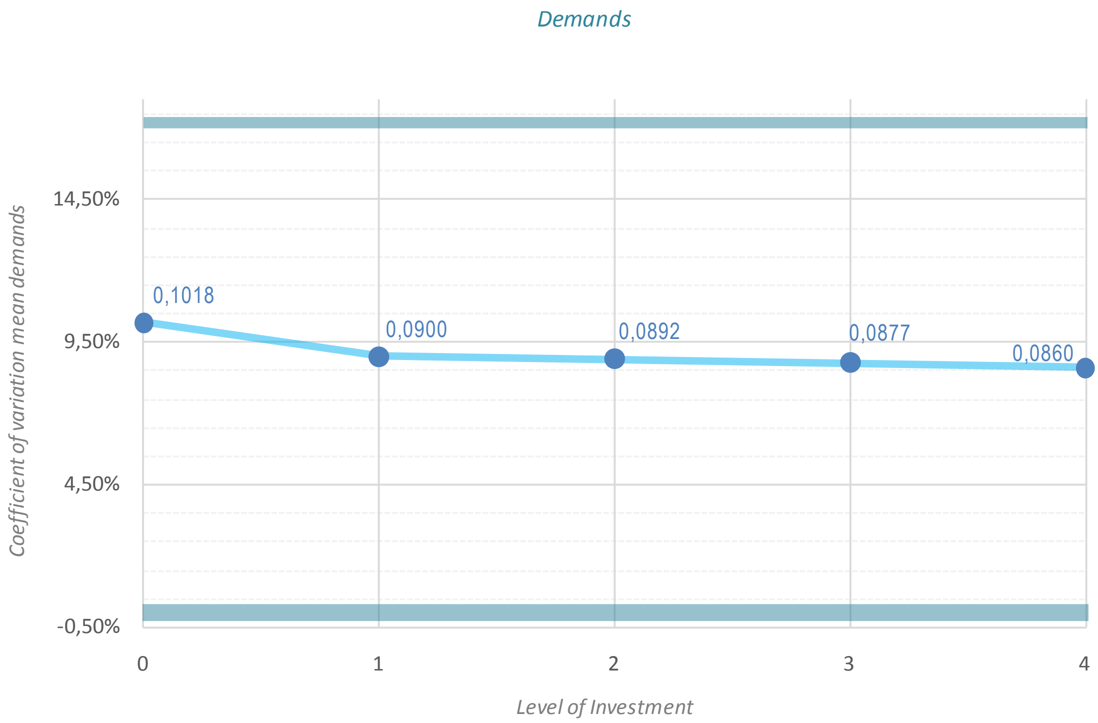
In turn, the evolution of the aggregate uncertainties for demands, flows in strategic network pipes and piezometric levels are shown in Figures 4, 5 and 6 respectively, in which the corresponding upper limits (defined by the baseline scenario that considers the existence of pseudo-measurements alone) and lower limits (for the ideal scenario of nil uncertainties) are also shown. It should be pointed out in these results that the uncertainty, with the current measurement configuration, is already far below the upper limit considered. In turn, in outline, the effect of the investment is positive, from the point of view of reducing uncertainty.

Singling out the case of the demands, as seen in Figure 4, a sharp drop occurs in coefficient of variation terms at the first investment step, which reduces progressively with additional investments, i.e. system saturation occurs progressively. Note that this first step includes the placement of the last demand flow meters at the outlets to the sectors, hence the reduction in subsequent investments occurs on indirect measurements and device redundancy.

The same happens with the flows in pipes, as shown in Figure 5, as a sharp drop also occurs in coefficient of variation terms at the first investment step, which reduces progressively with additional investments.

In turn, a first step effect is also seen in the uncertainty of piezometric heights, with a very slight reduction in uncertainty occurring with additional investments, as shown in Figure 6.

FIGURE 4. EVOLUTION OF AGGREGATE UNCERTAINTY IN DEMANDS WITH DIFFERENT LEVELS OF INVESTMENT



INDEX

FIGURE 5. **EVOLUTION OF AGGREGATE UNCERTAINTY IN FLOWS IN PIPES WITH DIFFERENT LEVELS OF INVESTMENT**

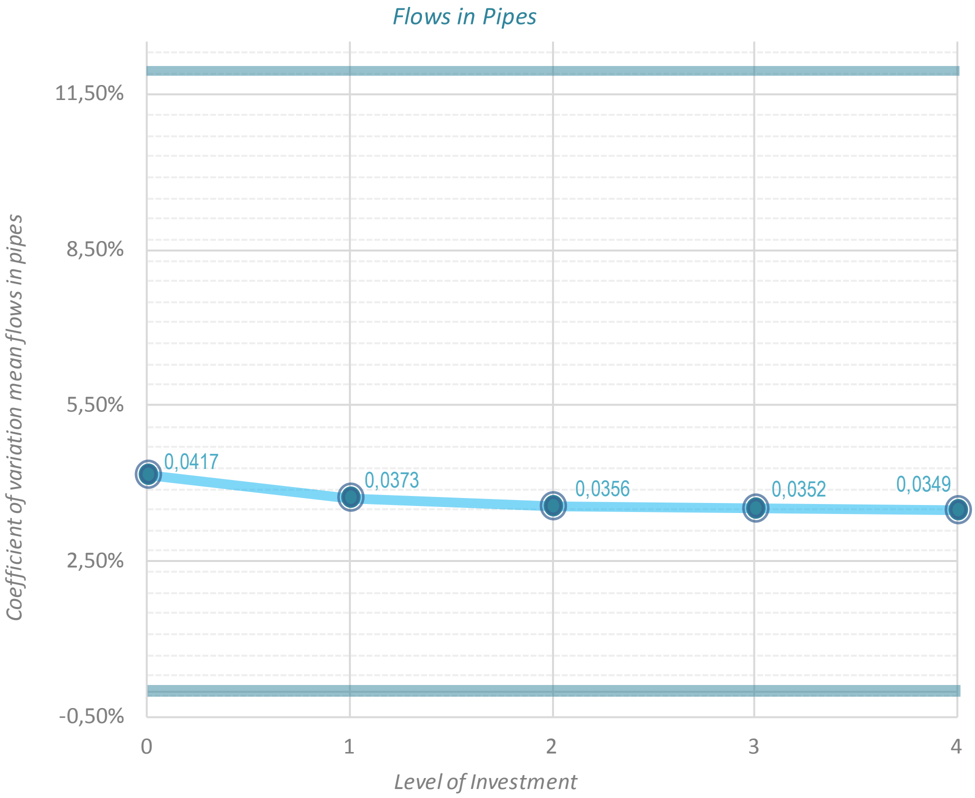
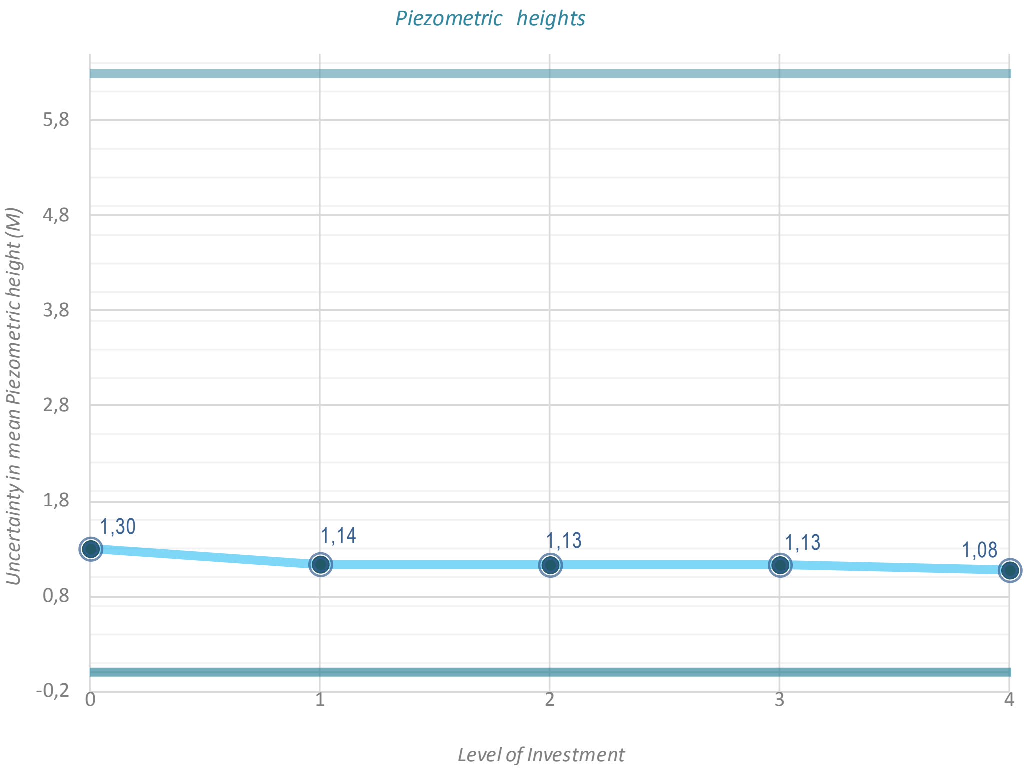


FIGURE 6. **EVOLUTION OF AGGREGATE UNCERTAINTY IN PIEZOMETRIC HEIGHTS WITH DIFFERENT LEVELS OF INVESTMENT**



For the rest of the parameters corresponding to less numerous elements, a drop in mean uncertainty is also seen as the level of investment increases. However, there are situations in which the uncertainty for higher investments is lower than that obtained with lower investments. For example, the coefficient of variation for the flow in the valves goes from 1.64% at the third level of investment to 1.66% at the fourth. This behaviour is explained because the average is calculated with a different number of elements, as the percentage of elements in which the uncertainty calculation is valid and is not affected by numerical errors (see the point entitled “Adaptation of the methodology to the strategic network”) changes for the different levels of investment considered, as shown in Table 12.

TABLE 12. PERCENTAGE OF VALIDITY FOR UNCERTAINTY CALCULATION OF NUMERIC ERRORS

Aggregate Results	*Level of investment*				
	Current	*1*	*2*	*3*	*4*
Coefficient of Variation for demand	98.82%	96.12%	97.05%	97.23%	96.91%
Uncertainty in piezometric height	99.21%	98.75%	98.93%	99.00%	98.84%
Coefficient of Variation for flow in pipes	99.36%	98.90%	99.08%	99.17%	99.02%
Coefficient of Variation for flow in pumps	96.91%	95.88%	95.88%	95.88%	94.85%
Coefficient of Variation for flow in valves	99.25%	99.25%	99.25%	98.50%	99.25%
Coefficient of Variation for demand in sectors	96.33%	77.03%	78.79%	82.14%	81.98%
Uncertainty in flow between sectors	96.37%	78.66%	84.80%	84.99%	84.32%

The results shown to this point are the absolute values obtained in terms of uncertainty and/or coefficients of variation for the aggregate parameters presented at the start of this section. However, the relative results, with respect to the minimum and maximum precision that could be achieved in the network, provide more information on what the investment means or brings. In this regard, Table 13 shows the ratios of the aggregate results for uncertainty, at the various levels of investment, with respect to the reference scale. To generate this, the results in Table 11 were scaled considering that the maximum possible uncertainty corresponds to the baseline scenario (in which there are only pseudo-measurements), which is assigned the value 0 as the worst case, and that the ideal situation is that of nil uncertainties, the theoretical scenario that is assigned a score of 10.

TABLE 13. **RATIO OF THE AGGREGATE RESULTS FOR UNCERTAINTY, AT THE VARIOUS LEVELS OF INVESTMENT, WITH RESPECT TO THE REFERENCE SCALE**

Aggregate Results (0 – 10)	*Level of investment*				
	Current	*1*	*2*	*3*	*4*
Coefficient of Variation for demand	4.09	4.78	4.83	4.91	5.01
Uncertainty in piezometric height	7.93	8.19	8.20	8.20	8.29
Coefficient of Variation for flow in pipes	6.50	6.87	7.01	7.05	7.08
Coefficient of Variation for flow in pumps	6.78	7.05	7.05	7.05	7.17
Coefficient of Variation for flow in valves	8.98	9.09	9.11	9.13	9.12
Coefficient of Variation for demand in sectors	2.48	4.45	4.18	4.20	4.18
Uncertainty in flow between sectors	2.90	6.83	6.38	6.38	6.39

0->Worst case level of uncertainty associated with the baseline case
10->Ideal level of uncertainty (nil)

These results show that the uncertainty of certain variables is already relatively good with the current measurement configuration, as in the case of piezometric levels at the strategic network nodes and the flow through the valves in the system, which do not improve significantly on increasing investment. On the contrary, the uncertainty for flow between sectors starts with the worst case, but the progressive addition of devices gives a significant improvement. In this regard, Table 14 shows the number of additional devices installed with respect to the current configuration for the various levels of investment. It can be seen in it that the number of flow meters and pressure meters increases progressively.

TABLE 14. **NUMBER AND TYPE OF DEVICES PROPOSED FOR EACH LEVEL OF INVESTMENT**

Type of devices	*Level of investment*				
	Current	*1*	*2*	*3*	*4*
Flow meters	-	26	58	84	92
Pressure meters	-	17	25	42	72

These results may appear contradictory initially, because, as stated above, the piezometric levels are already associated with relatively low uncertainties in the initial situation. However, the fact that the strategy for proposing monitoring improvements returns a higher number of pressure sensors with each investment is justified by the low cost associated with these devices and due to the fact of considering that whenever a flow meter is installed at a new location, which involves an additional cost in civil works to install it physically, a pressure sensor is placed as well. This strategy involves an appreciably lower cost than the price of the typical electromagnetic flow meter and the civil works, achieving an improvement in system uncertainty while investing the smallest possible amount of money.

Finally, it should be stressed that all the flow meters installed with investment level 1 are also installed at investment level 2, also including at this second level, 11 of the 17 pressure sensors installed at investment level 1. This trend is seen at all the levels, as, for example, 83 of the 84 flow meters selected at level 3 are also placed at investment level 4 and 31 of the 42 pressure sensors installed at the previous level are also placed at the maximum investment level. These results show that the objective function used to set out the optimisation problem making it possible to detect the most appropriate locations for placing metering devices given a certain level of investment not only reduces system uncertainty, but also tends towards the same solution to the problem, as the locations proposed for each investment package contain the majority of the locations obtained for smaller investment packages. This would enable the methodology to be used sequentially, so as not to stray too far from the overall optimum solution that could be achieved if the investment strategy was not stepped.

9. Conclusions

INDEX

This chapter covers the main conclusions drawn in this research in terms of identifying observable islands, quantifying observation uncertainty and locating additional metering devices. Furthermore, this section ends with an evaluation of the viability of implementing state estimation techniques to large supply networks, such as the case of the network managed by Canal.

9.1. IDENTIFICATION OF OBSERVABLE ISLANDS

The identification of observable islands is the first phase of the observability analysis presented in this work and makes it possible to identify the zones in which reliable results can be obtained in a subsequent state estimation process, as at least the algebraic relationships are sufficient to obtain a result. The selection of an appropriate method for implementing the identification of observable islands in the strategic network was performed by means of a multi-criteria analysis in which various aspects (resulting information, computational cost, agility for combination with state estimation techniques, etc.) were evaluated for different methods existing in the literature, adapted to water distribution systems. As such, after evaluating the null space, optimisation, algebraic, binary and stochastic methods, the **algebraic method** was found to be the most competitive, as this strategy makes it possible to evaluate the observability of all the network hydraulic variables and, consequently, identify the observable islands, as well as those that can be updated in the case of an isolated modification.

Thanks to the CHYPRE system Canal has for estimating consumptions in sectors and to the knowledge of the normal network operating conditions, the system, with these pseudo-measurements for consumptions and those for levels in tanks, is inside an observable island in its entirety. However, the interest in identifying observable elements lies especially in analysing whether modifications could be detected with respect to the reference state. In this regard, individual analysis were performed to identify the pumps and valves that would remain inside an observable island if the characteristic curve with which the element operates is placed in doubt, the diversion flow rates that could be caused if an initially closed valve is opened and the ability to detect breaks. The results obtained showed that between 53% and 81% of the operating state or point of the pumps and valves would still be in an observable island if an isolated variation occurred in the network, 42% of the diverted flow rates could be maintained in an observable island and algebraic relationships to detect and quantify breaks could only be established for 9% of the transit nodes that make up the strategic network.

This shows that the network is relatively resilient in observability terms, i.e. it has the capacity to adapt to the changes that may occur without an excessive loss of information, at least in terms of detecting changes in the network topology, as the detection of breaks in the system requires a specific study. This justifies the approach of locating devices to improve monitoring in terms of reducing uncertainty, as the behaviour of the system in terms of making it possible to identify topological changes is good. It is true to say that the analysis performed only considers the occurrence of individual variations with respect to the reference state, but large changes from one moment to the next are not to be expected, rather operating changes in the network would be expected to be more progressive. In this regard, it needs to be borne in mind that the final aim of state estimation is real-time system monitoring, hence only changes between successive time steps need to be taken into account.

9.2. QUANTIFICATION OF OBSERVATION UNCERTAINTY

Quantifying observation uncertainty represents a fundamental part of the observability study performed in this research, as it makes it possible to identify the most or least credible zones in the distribution system with a view to implementing a subsequent state estimation process. Its implementation in large networks calls for the quantification of the uncertainty inherent in a constrained state estimation problem, on which sophisticated methods have been adapted to reduce the existing order of magnitude differences and so minimise the associated numerical errors.

This methodology made it possible to tackle the calculation of the uncertainties corresponding to maximum, mean and minimum demand scenarios for the specific case of the strategic network managed by Canal, under normal operating conditions. The analysis performed made it possible to detect that the observation uncertainties in terms of pressure, flow rate and consumptions are reasonable, although it is true that these are somewhat impaired in the network zone corresponding to the distribution of flow between sectors and more so as system demand reduces. This fact could limit the ability to detect abnormalities in a subsequent state estimation process, with quantification of uncertainty being a tool that makes it possible to understand the certainty of the estimate.

However, it is worth stressing that on comparing the uncertainty of the variables associated with the current measurement configuration with the worst case scenario of the existence of pseudo-measurements (for demand and tank levels) alone, corresponding to a score of 0, and the ideal scenario in which the system variables have nil uncertainty, corresponding to a score of 10, a dimensionless ratio varying between 2.48 (demand in sectors) and 8.98 (flow through valves) is obtained. In this regard, it needs to be highlighted that all the strategic network variables have a score of over 5, while the scores are rather worse in the sectors and the connections between them. These results show that there is already an acceptable level of precision, but even so strategic locations are proposed for adding metering devices.

9.3. STRATEGIC LOCATION OF ADDITIONAL METERING DEVICES

The siting of additional metering devices was performed in this work with the aim of reducing observation uncertainty, as observability (or inclusion within an observable island) proved to be achievable for the entire network and relatively resilient according to the previous results. The adaptation of this methodology to a network of the size of that managed by Canal makes it necessary to carry out an optimisation process that considers a series of shortlisted locations, from among which the most advisable device configuration is selected.

The installation of additional metering devices was proposed considering four successive independent investment packages. This gives progressive improvement in system uncertainty as a result, achieving average scores between 0 and 10 (0 for the worst case scenario associated with the baseline case and 10 for the ideal scenario of nil uncertainty) of 6.75 for investment level 1 and 7.04 for investment level 4.

These results show a clear improvement over the current state of the network, although it is true that the lowest values continue to be recorded for the sectors and the connections between them. They also show that system saturation tends to occur, and from the second level of investment no significant improvements at the aggregate level are seen with increasing investment.

9.4. VIABILITY OF STATE ESTIMATION

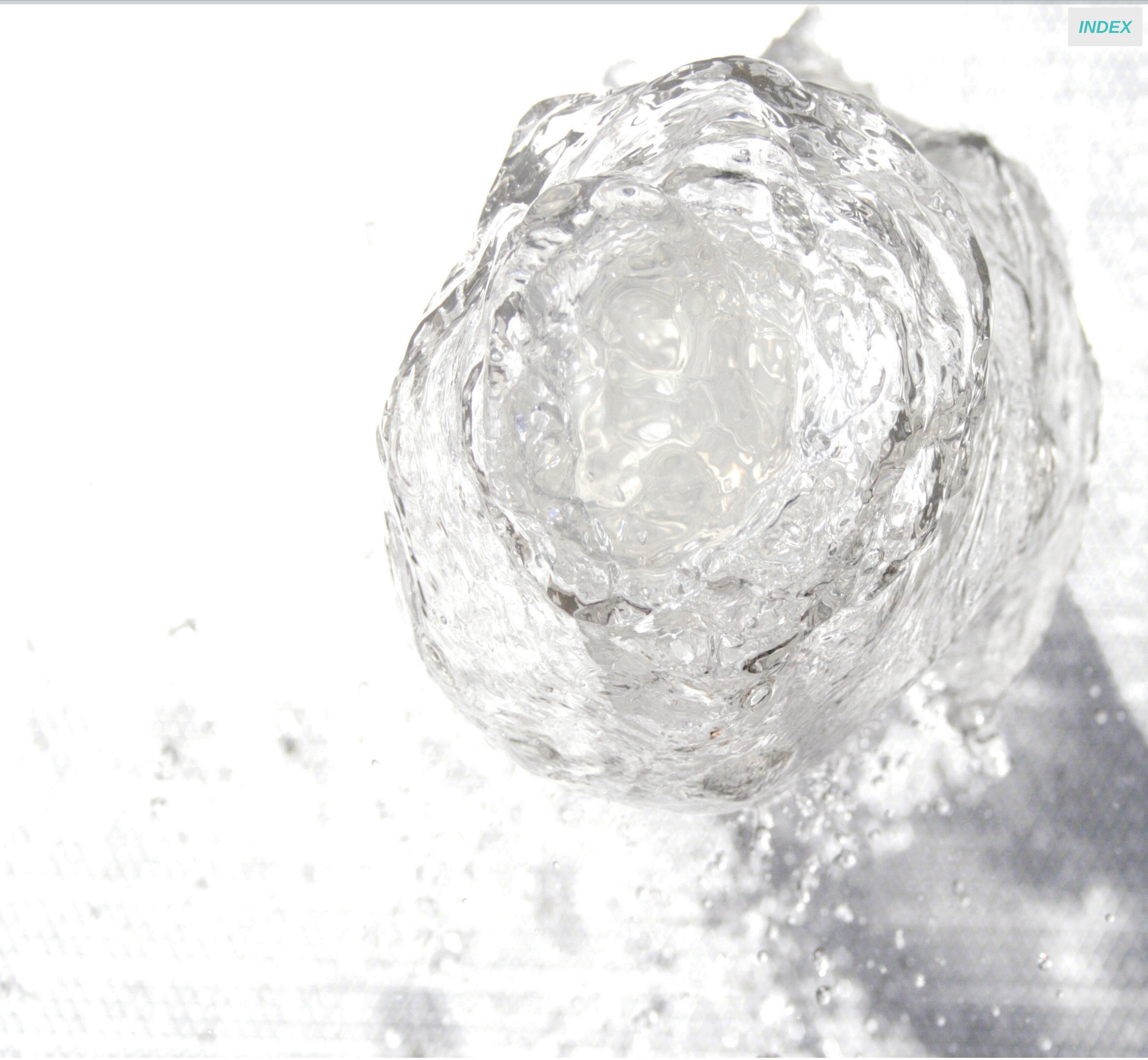
Having characterised the level of observability inherent in the network managed by Canal de Isabel II, it is possible to evaluate the viability of implementing state estimation techniques for the system. In this regard, the good level of observability obtained for normal operating conditions ensures that it would be possible to monitor the behaviour of a significant proportion of the network in real time, even in the case of a number of changes in network topology (unnotified operations), although it is true that the isolated loss of system information (metering devices) could reduce the number of types of operations that cannot be identified and reduce the precision of the estimate determining the statistical significance of the estimate. The level of precision the system state estimation would achieve is maintained over a large part of the network and is clearly superior to that contributed by the baseline situation that only considers the pseudo-measurements, which shows that dependence on these in state estimation is limited only to certain zones of sectors under construction, still not closed and, above all, to connections between sectors where there are no flow meters. As such, the observability analysis performed is a fundamental tool which needs to be incorporated into the state estimation process for continuing evaluation of it, as it makes it possible to identify in which zones results are being obtained and to what extent they are reliable.

The observability study performed on Canal's network clearly shows the high potential that remote metering systems implemented in sectorised networks offer for understanding the state of the system in large transport networks, making it possible to extend the ability to observe and identify particular network conditions beyond the directly measured variables. It is shown, therefore, that hydraulic state estimation, as a subsequent dynamic tool, has a great potential capacity for monitoring the network state, giving early detection of significant incidents that have occurred and that have not been notified, which will make it possible to offer a high value service for decision making in operation.

On the other hand, the methodologies developed and implemented during this project, which made it possible to analyse the uncertainty in the possible observations of a state estimator for a network the size of the network managed by Canal, show the technical viability of implementing a state estimator.

APPENDICES

INDEX



APPENDIX 1. GLOSSARY OF TERMS

This section defines and explains some terms used throughout the document. In particular, it details the way they should be interpreted in the context of this work.

CHYPRE

An application that enables estimation of consumption in sectors and the associated uncertainty in the network managed by Canal. See Carrasco and García (2011) for further detail.

DMA

District Metered Area. Referring to each of the sectors defined in a water distribution network, which are those areas where the inputs and outputs are limited and are typically measured.

DWTP

Drinking Water Treatment Plant

FOSM

First Order Second Moment. A technique enabling the uncertainty of some variables to be propagated to others on the basis of a first-order Taylor series.

GIS

Geographic Information System. Referring to the set of tools comprising various components to improve the storage, handling and analysis of large amounts of information.

ICT

Information and Communications Technology. Referring to the group of information technologies deployed currently to improve the monitoring of various processes.

Observability

A property of the variables involved in a hydraulic problem. A variable is said to be observable when there are enough measurements available to establish algebraic relationships which make it possible to infer the value of the variable. A network is said to be observable when all its variables are.

Observable island

A drinking water supply network zone where the variables are observable. This makes it possible to identify sub-zones in which state estimation would provide reliable results without observability being achieved in the whole network.

Observation uncertainty

The standard deviation inherent to the observation of the network variables, which can refer to flows, consumptions and piezometric levels. It can be estimated based on the characteristics of the flow network and the available measurements, as a second phase of the so-called observability analysis.

SCADA

Supervisory Control and Data Acquisition. A platform enabling data collection in real time from the existing measurement devices in a system, in this case a supply network.

State estimation

A procedure or technique that makes it possible to infer the flow conditions of a water distribution network by means of the least squares fit of the error between the system state variables and the existing measurements. Throughout this work, pseudo-static state estimation is referred to, independent of time.

Strategic network

Understood as such is the network bringing together the main arteries of the network managed by Canal, which allows water to be transported to the sectors that make up the Comunidad de Madrid drinking water supply network.

SVD

Singular Value Decomposition. A technique enabling the factorisation of matrices, making them easier to process.

SOI

Stochastic Observability Index. An index that makes it possible to evaluate the general observability of a water distribution system in terms of piezometric levels or flow rates.

SOWI

Stochastic Observability Weighted Index. An index that makes it possible to evaluate the overall observability of a water distribution system by weighting the uncertainties of the flows themselves circulating through each pipe by the flow value itself.

APPENDIX 2. REFERENCES

Carrasco, A. y García, L. 2011

Desarrollo de un Sistema de validación, estimación y predicción de consumos horarios por sectores para la red de distribución de Canal de Isabel II. Cuaderno de I+D+i nº13. Madrid, Canal de Isabel II.

Castillo, E., Conejo, A.J., Pruneda, R.E. and Solares, C. 2005

State estimation observability based on the *null space* of the measurement Jacobian matrix. *IEEE Transactions Power Systems*, 20(3), 1656-1658.

Díaz, S., González, J. and Mínguez, R. 2015

Observability analysis in water transport networks: algebraic approach. *Journal of Water Resources Planning and Management*, doi: 10.1061/(ASCE)WR.1943-5452.0000621.

Díaz, S., González, J. and Mínguez, R. 2016 a

Uncertainty evaluation for constrained state estimation in water distribution systems. *Journal of Water Resources Planning and Management,* doi: 10.1061/(ASCE) WR.1943-5452.0000718.

Díaz, S., Mínguez, R. y González, J. 2016 b

Aproximación estocástica al análisis de observabilidad en redes de abastecimiento de agua. *Revista Ingeniería del Agua*, 20(3), doi: 10.4995/la.2016.4625.

Habiballah, I.O. and Irving, M.R. 2001

Observability analysis for state estimation using linear programming. *IEE Proc. – Gener. Transm. Distrib.*, 148(2), 142-145.

Pruneda, R.E., Solares, C., Conejo, A.J. and Castillo, E. 2010

An efficient algebraic approach to observability analysis in state estimation. *Electric Power Systems Research*, 80(3), 277-286.

Solares, C., Conejo, A.J., Castillo, E. and Pruneda, R.E. 2009

Binary-arithmetic approach to observability checking in state estimation. *IET Gener. Transm. Distrib.*, 3(4), 336-345.

APPENDIX 3. INDEX OF FIGURES

APPENDIX 4. INDEX OF TABLES

Santa Engracia, 125. 28003 Madrid
www.canaldeisabelsegunda.es