Design of a monitoring network: from macro to micro design

Lepot, Mathieu; Kapelan, Zoran; Clemens, Francois H.L.R.

DOI
10.2166/9781789060119_0155

Publication date
2021

Document Version
Final published version

Published in
Metrology in Urban Drainage and Stormwater Management

Chapter 6

Design of a monitoring network: from macro to micro design

Mathieu Lepot[1,2], Zoran Kapelan[1,3] and Francois H. L. R. Clemens-Meyer[1,4,5]

[1]*Delft University of Technology, Water Management Department, Delft, The Netherlands*
[2]*Un poids une mesure, Lyon, France*
[3]*University of Exeter, College of Engineering, Mathematics and Physical Sciences, Exeter, United Kingdom*
[4]*Deltares, Unit Hydraulic Engineering, Dept. Experimental Facility Support, Delft, The Netherlands*
[5]*Norwegian University of Science & Technology, Faculty of Engineering, Dept. Civil & Environmental Engineering, Trondheim, Norway*

ABSTRACT

Designing a monitoring network or a measuring set-up or a monitoring station is a typical (multidisciplinary) engineering enterprise: a range of potentially conflicting demands (technical, financial and managerial) and limitations (e.g. availability of resources, skilled personnel, regulations) have to be respected. This chapter addresses the design aspects on both the macro scale (a monitoring network) and on the micro scale. The macro scale addresses what to measure, where to measure, how frequently to measure and the applications of models in the design process. On the micro scale issues with safety, accessibility and practical limitations are discussed. This chapter has close links with virtually all other chapters in this book and a comprehensive set of literature references is supplied to allow the interested reader to broaden his/her knowledge on the subject.

Keywords: Micro design, monitoring, network design, practical aspects.

doi: 10.2166/9781789060119_0155

SYMBOLS

C	weir coefficient ($m^{1/2}/s$)
f_N	Nyqvist frequency (Hz)
$f(n)$	inverse Fourier transform element 'j'
$F(k)$	Fourier transform for frequency k
h	water level (m)
i	imaginary unit, integer value (-)
I	amount of information (-)
j	index (-)
J	Jacobian matrix (various)
k	integer value (-)
k_n	Nikuradse's equivalent sand roughness (m)
loc	as exponent: discrete monitoring location index
m	number of measurement locations (-)
mse	mean squared error (various)
M	selected number of monitoring locations among N possibilities (-)
n	integer value, number parameter used for evaluating the Jacobian matrix (-)
N	number of monitoring locations (-)
N_e	number of elements in a time series (-)
$p_1, p_2,.., p_n$	parameters of a model (various)
P	power of spectrum
Q	discharge (m^3/s)
$r(T)$	autocorrelation function for the time shift T (-)
SNR	signal to noise ratio (-)
t	time (s)
t	as index: discrete time index (-)
T	time shift for which the autocorrelation of the process $x(t)$ is estimated (s)
$x(t)$	a time varying process (various)
α	weighing factor (-)
β	weighing factor (-)
γ	weighing factor (-)
Δt	sampling time interval (s)
μ_x	mean value of the process $x(t)$
$\rho(T)$	normalized autocorrelation function for the time shift T (-)
ρ_x	autocorrelation of the process $x(t)$ (-)
σ_{interp}	standard deviation of an estimated interpolated value (various)
σ_m	standard deviation of measurement uncertainty (various)
σ_n	standard deviation of the noise (various)
σ_x	standard deviation of the process $x(t)$ (various)
τ	a given value of time t between two successive times t_{j-1} and t_j (s)

Motivation anecdote 'Unexpected cycles' from Walcker *et al.* (2018)

In 2016, the OTHU (Field Observatory on Urban Hydrology) in Lyon, France set up the second generation of their monitoring stations. After a decade of experience, trials and feedbacks, new monitoring stations have been designed, built, and implemented (see also Section 6.3.5). Figure 6.0 (a) shows the discharges measured at the inlet of a stormwater retention tank with the old (red) and the new (blue) data acquisition methods.

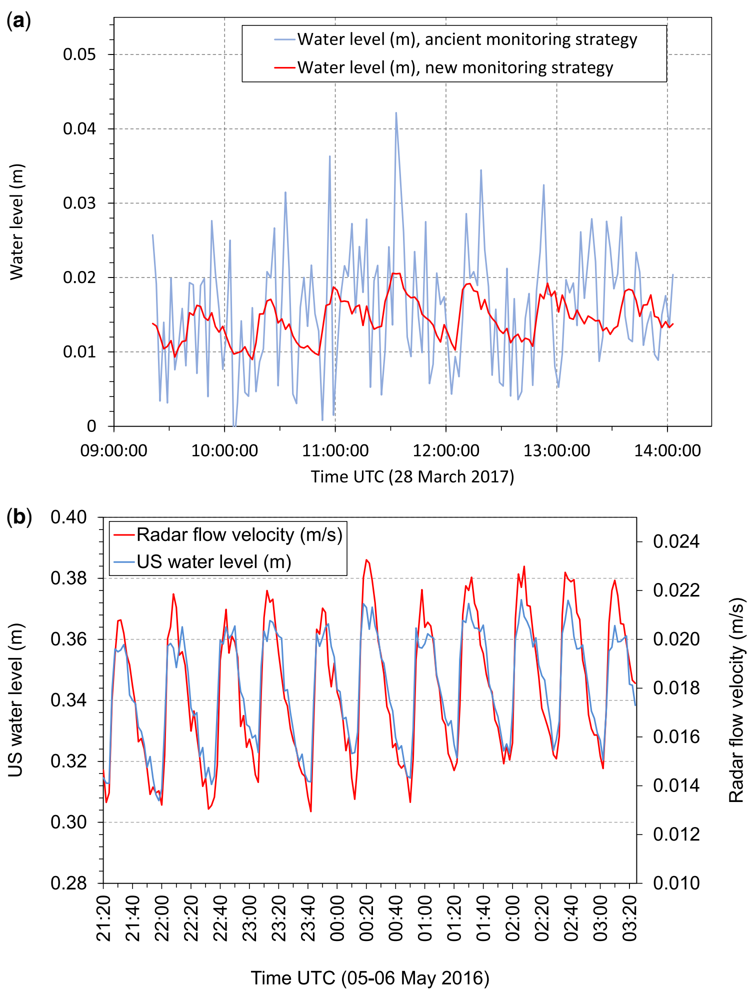

Figure 6.0 (a) water level – red line: old monitoring system, blue line: new monitoring system; (b) water level (blue line) and flow velocity (red line) measured with the new monitoring station. *Source*: Nicolas Walcker (INSA Lyon).

At first sight, the new station delivers less noisy data, mainly but not only due to the fact that the value recorded every two minutes is calculated from 120 values measured every second and no longer sampled as a single instantaneous value measured every two minutes. Looking at a few days within the time series, 30 min cycles (Figure 6.0(b)) become visible with those smooth data for different quantities: water level, flow velocity, discharge and, more surprisingly, pH and conductivity, even during rain events. Since those fluctuations occur with several sensors but not all, they cannot be explained by the data acquisition system: something happens within the sewer. Investigations have shown that water used for industrial processes is discharged into the stormwater pipe of the separate sewer system. The improved data acquisition method with on-line pre-processing of high frequency data made possible by a better monitoring system leads to new knowledge, and, in this case, to the identification of an illegal inflow within the sewer.

What can you learn if you (re)build a high-quality system? This chapter will give you few methods and tips to guide you in such projects.

Mathieu Lepot and Nicolas Walcker

6.1 INTRODUCTION

In this chapter the design of a monitoring set-up is discussed. On the one hand a distinction is made between set-ups for permanent or long-lasting use (e.g. weeks to years) and set-ups for occasional measurements. On the other hand, a distinction is made between theoretical and practical aspects of the design. Further the reader is referred to a range of matters strongly interwoven with design aspects that are discussed in other chapters. As this subject is a very comprehensive one, the reader is also supplied with entries to the literature for further study. Overall, when starting a monitoring project, a scheme as shown in Figure 6.1 is applied. A first omni-important step is to agree upon the data that needs to be gained from the monitoring set-up to fulfil the information need according to monitoring goal(s).

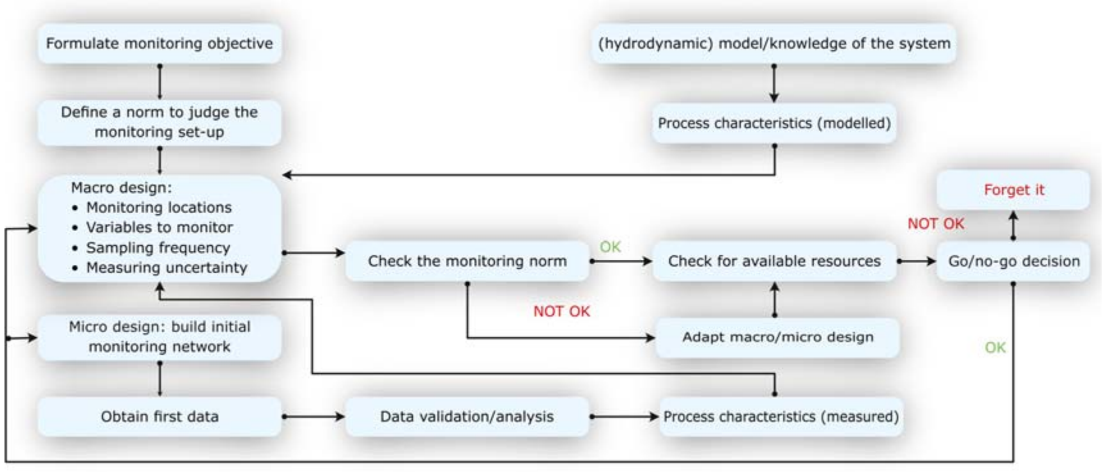

Figure 6.1 Basic flow diagram for designing monitoring networks, initially using a model or another source of *a priori* knowledge on the system to be monitored. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

From this information need, the following questions arise:

- What quantities(s) should be monitored?
- At which location(s) should monitoring take place (where and how many)?
- At what sampling frequency?
- What is the maximum allowed uncertainty?
- What should be the duration of the monitoring campaign?

When these questions are answered a first rough budget estimate can be drafted and a managerial decision whether or not to proceed with the project has to be made. Basically, the question '*Is the information obtained worth the investment and additional costs?*' has to be answered. There is no general 'recipe' on how to handle such a question, as in many cases the possible 'gain' or 'loss' cannot be entirely expressed in monetary terms (e.g. in cases of monitoring set-ups that are mandatory for legal purposes, as frequently included in permit conditions). In the case where it is decided to proceed with the monitoring project, the information yield has to be regularly checked against the original information need in order to allow for making adaptations of the monitoring set-up (e.g. reducing the number of locations, replacing sensors, reducing the sampling frequency, etc.). Such a check, or evaluation, should be done on a regular basis with an increased frequency at the beginning of the project, and for long-term projects at least once a year. An important driver here is the regular economical re-evaluation of the project: '*Do we still conclude that the information obtained is value for money?*'

In long-term projects the rationale for checking the information yield is found in the following considerations: (i) (small) adaptations in the structure and/or geometry of the system under study and (ii) a shift in the goals strived for. These situations result in a change in information obtained and a change in information need, respectively.

This implies, especially for long-term monitoring projects, that there is a strong need to obtain:

- Information on the quality of the data obtained (see Chapter 9 on data quality and validation).
- Insight into the costs involved for keeping the monitoring project operational (that is in a physical sense as well as in the sense that data gained are being processed into information and are actually used for the defined purpose).
- Actual data on adaptations made in the system under study, so as to enable linking system behaviour as recorded to the system's status (geometry, structure, operation) over the monitoring period. This is even more important when the goal of the monitoring project is to quantify the effect of some adaptation in the system under study, or when evaluating the effect of taking large-scale measures in a system to e.g. reduce the pollutant load into the environment.

Overall, it has to be realized that the design of a monitoring set-up is, certainly for long-term monitoring, not static. Further aspects to keep in mind when developing a design are the demands put on the organization responsible for keeping the system operational and budget claims implied by design choices made during the process.

The following general cases are distinguished:

(1) Monitoring as a basic need for the (asset) management of the urban drainage system.
(2) Monitoring for legal purposes (location specific, long- and short-term depending on the legal issue).
(3) Monitoring for regulation or environmental compliance (e.g. CSO (combined sewer overflow) spills, in- and exfiltration, wrong connections).
(4) Control purposes (location specific, long term).
(5) Decision support for a redesign of an existing system.
(6) Model calibration.

It is largely stated that cases 1 to 4 (and to a lesser extent case 5) are normally straightforward in terms of what and where to monitor and for what duration, the design effort is then concentrated on the micro design.

In theory one could argue that when taking 'Model calibration' (case 6) as a goal one could, once the calibration is proven successful, use the model as a surrogate reality with which all other questions arising from 1 to 5 can be answered. This is true only up to a certain extent. The following remarks have to be made in this respect:

- Even after calibration, significant uncertainties in the model parameters (and hence the model predictions) remain, limiting its effective applicability.
- Unnoticed, significant deviations with the 'real world' may exist as not all relevant processes are correctly incorporated in the present generation of models (see e.g. Tscheikner-Gratl *et al.,* 2019), limiting again its effective use.
- Regular re-calibration is needed for the same reasons that a monitoring set-up for long term monitoring has to be evaluated on a regular basis. With respect to case 5, monitoring a system in relation to a planned redesign or enhancement of functionality, sets the same demands on the monitoring design as model calibration and is not separately discussed here any further.
- Application of the model parameters obtained beyond the domain for which they are determined (i.e. the range of loads for which the parameters are calibrated), may result in biased results.

Key messages on design of monitoring networks and stations

- KM 6.1: *Macro ~ Micro design* – 'Where and what to measure' questions are answered by the macro design approach. The micro design answers the question of 'How'.
- KM 6.2: *Mathematics vs. experts* – Expert designs are still often more reliable than the ones determined by mathematical methods.
- KM 6.3: *Teamwork* – Both designs require teamwork: never neglect the feedback from all the parties involved in the project.
- KM 6.4: *1, 2 and 3* – If you are beginner and/or too optimistic, please keep in mind the rule of '1, 2 and 3': 1 rule, the budget will be 2 times as expensive than the first expectation and the forecasted duration should be multiplied by 3.

6.2 MACRO DESIGN

Macro design of a monitoring network encompasses the choice of the number of monitoring locations, what to measure, how frequently and with what quality in terms of uncertainty in the monitoring results and data. By definition, this is a cyclic process. The initial design is based on the knowledge available on the system at hand, while the monitoring system is meant to extend and deepen the knowledge on the system. This implies that after obtaining and interpreting the data, they may hold clues for further refinement of the monitoring system. When developing the macro design it has to be realized that choices made here may affect the 'margin of freedom' in the micro design. For example, when setting up criteria for eliminating or identifying potential locations, and not taking into account the availability of certain services (data communication or power infrastructure), this limits the choice for the type of power supply and data storage and transfer methods.

6.2.1 General

For a wide range of monitoring objectives, the monitoring location(s) is(are) unambiguously defined, e.g.:

- What is the discharged CSO volume at this CSO structure?
- Does flooding occur on a specific location?
- What is the performance of this pumping station?
- What is the mass balance in a given catchment?

In other words, often the formulation of the monitoring goal explicitly defines where and, to a certain extent, what to measure. There are however cases in which this is not entirely clear, when e.g. the following monitoring goals are formulated:

- We need monitoring to calibrate a model.
- The hydraulic impact of the urban drainage system on the river must be quantified using monitoring data.

Choosing monitoring locations for such goals requires some prior information/knowledge. The system(s) should be known in some detail (structure and geometry, details on land use, connected surfaces, information on ground water levels, locations where flooding occurred, citizen's complaints). There should be means to get a preliminary impression of the system's response to loads (i.e. storm events and/or wastewater discharges). Such a preliminary impression may be supplied by using a model simulation, although simpler data can be useful as well. When translating the choice of monitoring location(s) into an engineering question, the following task is: *'Given the available budget and the monitoring goals identified, find the minimum number of monitoring locations and their actual locations in the system.'*

In some cases, the budget will be insufficient to achieve the goals set. In such a case one either needs to raise the budget (political/managerial decision) or try to find cheaper methods to achieve the goals. If none of these options is applicable, one has to abandon the idea of monitoring altogether. However obvious the latter conclusion may seem to be, in practice parties often implicitly proceed in such cases with reduced ambitions and/or poorly designed monitoring set-ups, which in most cases ends in a disappointment. It is this type of situation that erodes away the political/managerial support for monitoring campaigns.

Hereafter an approach for choosing monitoring locations based on expert judgement is discussed as well as the added value (computer) models may bring in the design process.

6.2.2 Choosing locations as a combinatorial problem

Now, let us assume that it is possible to express the monitoring goals as an 'amount of information I' to be obtained from N monitoring locations, where N is defined by the available budget and the choices made for sensors, data handling, etc.

Suppose that there are only M locations that allow for monitoring, which in many cases are in manholes, as they are the main entrances for most underground urban drainage systems. Typically, M is much larger than N. This results in a discrete optimization problem:

> *'Find the combination of N manholes, out of a population of M possibilities, that maximizes the information obtainable from these manholes and check whether this is sufficient to achieve the monitoring goal(s) set.'*

In practice, the number of possible solutions is immense: for example, even for a relatively small case with $N = 28$ and $M = 210$, the number of possible solutions is approximately 5×10^{34}. Since it is a discrete problem, popular optimization algorithms like e.g. Levenberg-Marquardt, Simplex Method, etc.

are not applicable in a straightforward manner. Apart from that, this type of optimization problem is suspected to be of NP (hard) nature (NP stands for 'Non-deterministic Polynomial-time' and refers to a class of problems for which there is no algorithm known that can solve such a problem in a predictable number of operations. A famous optimization problem in this category is the Travelling Salesman Problem and algorithms to solve it, e.g. in Diaz-Delgadillo *et al.* (2016)). This implies that there is likely no algorithm that will give the optimal solution within a limited and predictable calculational effort. Therefore, for all practical purposes, one has to settle for a 'good' solution rather than striving for the 'best'. These are the main reasons why optimizations are usually carried out with either genetic algorithms or simulated annealing algorithms (e.g. Boomgaard *et al.*, 2001; Ruiz-Cardenas *et al.*, 2010). Obviously, a very effective manner to reduce the search space is to reduce the number of possibilities to choose from, i.e. eliminate all locations that cannot be considered as a monitoring location based on practical considerations (like e.g. accessibility, safety, etc.). In the literature (e.g. Clemens, 2001; Henckens & Clemens, 2004; Thompson *et al.*, 2011) different theoretical problem formulations exist with associated sensor network design criteria and decisions to be taken. All these approaches and theoretical frameworks are ultimately based on an extended analysis of sensitivity to parameter variation of a model. Much of the theory and algorithms are developed for pressurized systems (e.g. water supply networks), and transferring these algorithms for application to urban drainage systems (in which free surface flow and pressurized flow can both occur and holding transitions between the two modes over time) is not straightforward. The methods presented so far prove to be hard to implement and are not convincingly better by any metric (costs, quality of data obtained or reliability of monitoring results) than a design obtained by expert judgement. Although the theory of optimization of a monitoring network is an interesting topic for further development, its practical applicability is at present judged to be very difficult and is therefore not discussed here in detail. References to the existing literature on the subject are provided.

6.2.3 Considerations in choosing locations

Sensor network design involves deciding on what, where and under what conditions to measure. This approach relies on some prior knowledge of the system (either expert knowledge, or existing observations or a model mimicking the system behaviour) along with related knowledge of sensors, their costs, accuracy, and potential installation, operation and maintenance issues, etc. Whilst expert knowledge can be quite valuable, especially in specific circumstances, the ultimate decisions (i.e. the sensor network design to be chosen) tend to be subjective in nature, for obvious reasons. It has to be appreciated, however, that designs based on mathematical algorithms (as briefly mentioned in the preceding paragraph), e.g. to optimize information content, do not automatically acknowledge the added value of redundant information for validation purposes or any other practical circumstance other than the elimination of locations for practical considerations. This statement should be balanced by several facts: (i) actual goals of the sensor network should be clearly defined, (ii) future goals must be considered, (iii) the data on facilities and their accessibility must be implemented correctly, and (iv) evolutions (extensions of the system, rehabilitation activities or decommissioning of elements) of the catchment must be considered.

Even though choosing the 'optimal' set of monitoring locations is, to a certain extent, a subjective process, it is still mainly based on sound engineering criteria and it involves the following practical aspects:

- Purpose of the sensor network. Experts are driven by solving a specific sensor network design problem, e.g. designing a sewer network for a specific purpose. Having said this, quite often, they tend to think beyond that purpose, i.e. tending to think about other possible future applications and, generally, what might change in the future that will impact on their decision where to locate and what sensors to choose. For example, an expert may be aware of the fact that a town and its urban drainage system will expand in the easterly direction in the near future and, as a consequence, they may decide to locate a new flowmeter and level monitoring in that part of the town even though there is no real need to do so based on the current situation.

- Accessibility. Urban water systems are often comprised of large pipe networks that stretch over significant geographical areas, both urban and rural. As a consequence, some of the potential sensor locations, e.g. remote ones or locations in parking lots or in the proximity of heavy traffic, may not be easily accessible and an expert may choose not to install a sensor in such locations for these reasons.
- Safety. Working in the urban environment can be challenging, as one needs to account for a range of hazards, e.g. traffic, working in confined spaces, risk of contamination when working with polluted (waste)water, etc. Chapter 7 touches upon these issues in more detail. It is, however, efficient to take safety aspects into consideration early in the initial design.
- Availability of power supply. Most measuring devices require power to run. Some of them can run on batteries but there is always an issue of battery life which, in turn, is linked to operation costs. An alternative is to obtain the power from a nearby property but this often leads to accessibility issues. Experts are very aware of these issues which tend to impact their decisions on where to locate sensors and what sensors to use in the first place.
- Data communication. Many modern sensors communicate with a control room in a water utility or between themselves. This is usually done nowadays by means of wireless communication. There is no point in installing such a sensor in a location where the mobile network coverage is poor when relying on GSM (global system for mobile communication).
- Security. Most modern sensors are not cheap and some may be very expensive. Consequently, they need to be secured against theft/vandalism and this is easier to do in some locations than others. Moreover, modern sensors can potentially be remotely manipulated, e.g. via internet, and this may result in triggering undesirable actions in automated systems, e.g. at treatment works or systems that utilize real time control (RTC). This requires not just ensuring the physical protection of sensors but also that cyber security is ensured.
- Budget and availability of sensors. Practitioners are very much aware of budget constraints and limited availability of sensors (within their organization or otherwise), which has a major effect on decisions made regarding the sensor network design. Over recent years, cheap DIY (Do It Yourself) electronic systems (e.g. Arduino®) have been introduced. These systems allow laymen to put together a sensor system that is cheap and easy to obtain. Although no comprehensive objective data are available to date, the authors' experience hints at issues with reliability, and sensor quality in a broad sense, along with issues related to the operational conditions in the field in general. However, for small-scale trials or experiments, the application of DIY systems may prove to be a future game changer for monitoring the urban environment, as it comes at low costs and therefore allows for making errors or misjudgements in the design without substantial (financial) consequences. Recently, a protocol for testing low-cost water level monitoring was suggested which will allow sharing of experiences with such sensors on a common basis (Cherqui *et al.*, 2020).
- Make sure there is some overlap/redundancy of the expected recorded values to allow for consistency checks when validating the data.

Figure 6.2 shows an example of two subsystems connected at some point via manhole C. Suppose this is a combined sewer system, the subsystem connected to manhole A and the subsystem connected to manhole B

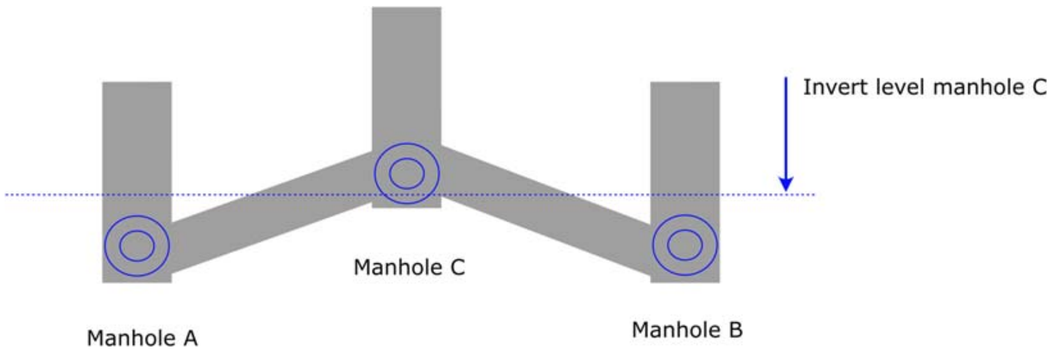

Figure 6.2 Example of a geometry that allows for overlap in readings for manholes A and B, when the water level reaches the threshold of the invert level at manhole C. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

will show a high correlation between monitored water levels at these nodes when the water level rises above the invert level indicated in the figure. During dry weather flow and moderate storms, this connection is not present. In the case of larger storm events, an occasional overlap in water levels is present allowing for cross-checking the readings obtained in manholes A and B.

The option of changing measurement locations after a first evaluation of the network information yield should be kept open, certainly when the system is meant for long-term monitoring. Modelling the system can be informative for the expert, especially if he/she is able to perform a sensitivity analysis with respect to model parameters. The model can assist the expert in identifying locations in which two or more sensors may show some overlap in the expected data.

Expert knowledge-based sensor network design is typically carried out according to the following general principles:

- Ensure good network coverage of the analysed urban water systems. Experts tend to distribute sensors relatively evenly through the analysed pipe network/geographical area. They know, intuitively, that good network coverage is important regardless of the intended sensor network use. This ensures, among other things, an effective everyday monitoring of the system and is especially important for detecting various events in pipe networks (e.g. collapses and blockages in a sewer network). In addition to distributing the sensors evenly in space, experts know that a number of sensors need to be distributed toward the pipe network edges. Otherwise substantial parts of the pipe network may end up not being observable at all. In dendritic systems (i.e. networks with very few loops and interconnections), prior reasoning may be a very effective manner to decide where to monitor.
- Install sensors at important and key system locations. Experts know, based on experience, intuition and a 'feeling' of the system they manage, that observing some of the key urban water system elements and structures (e.g. locations where flooding occurs, pumping stations, large sewers, outfalls, CSO constructions, etc.) is critical to ensure good system observability and ensure normal system operation.
- Use good quality and reliable sensors. Experts know, based on experience, that going for cheaper (hard- and software) solutions now is likely to result in additional efforts in, amongst others, maintenance and data validation during the project.
- Calibrate (Section 7.6) and maintain (Section 7.4) the sensors regularly. This is essential for the effective operation of the designed sensor network.

6.2.4 Example of using a model as a design aid

When a model is available, it can be used in identifying locations that provide information related to certain model parameters. Basically, a sensitivity analysis is performed on the response of the model (water level and/or discharge) when changing parameter values. When this is done in a systematic way, the potential of each location with respect to each individual model parameter can be obtained.

Consider a basic example of application, with two parameters: the hydraulic roughness k_n and the weir coefficient C. Three simulations are made using a hydrodynamic model, one with the parameter vector $[k_n, C]$, and two with parameter vectors $[k_n + \Delta k_n, C]$ and $[k_n, C + \Delta C]$. Thus, for each node (manhole) of the system, three hydrographs are obtained.

Based on these model results, the Jacobian matrix is built (Figure 6.3). For two parameters p_1 and p_2, the elements in the $m \times 2$ Jacobian matrix J (Equation (6.1)) are defined as:

$$J = \begin{bmatrix} \frac{dh_{t=1}^{loc=1}}{dp_1} & \cdots & \frac{dh_{t=1}^{loc=m}}{dp_2} \\ \vdots & & \vdots \\ \frac{dh_{t=n}^{loc=1}}{dp_1} & \cdots & \frac{dh_{t=n}^{loc=m}}{dp_2} \end{bmatrix} \approx \begin{bmatrix} \frac{\Delta h_{t=1}^{loc=1}}{\Delta p_1} & \cdots & \frac{\Delta h_{t=1}^{loc=m}}{\Delta p_2} \\ \vdots & & \vdots \\ \frac{\Delta h_{t=n}^{loc=1}}{\Delta p_1} & \cdots & \frac{\Delta h_{t=n}^{loc=m}}{\Delta p_2} \end{bmatrix} \tag{6.1}$$

where m is the number of locations where the water level is calculated, and n is the number of times the water level is calculated during a simulation.

The last term in Equation (6.1) is the partial difference approximation of the Jacobian J obtained using the model results. For each node, a Jacobian is obtained showing exactly for each time step the sensitivity of the water level to a variation in the parameters, each column corresponding to one parameter.

Figure 6.4 shows a very simple, artificial network (basically a hydraulic analogon of the well-known Wheatstone bridge electrical circuit). All nodes have a free surface area of 2 m^2, the conduits all have a

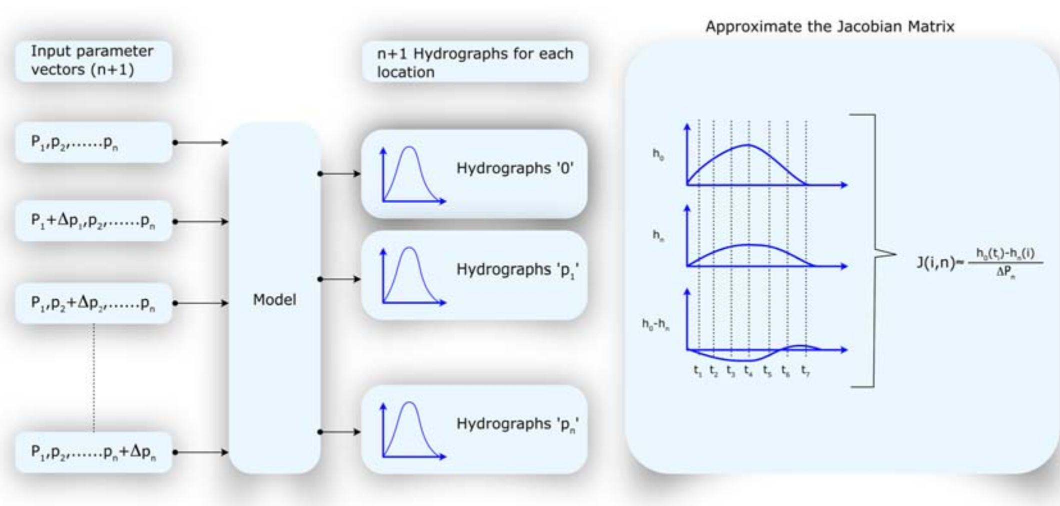

Figure 6.3 Scheme for the construction of the Jacobian matrix *J*. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

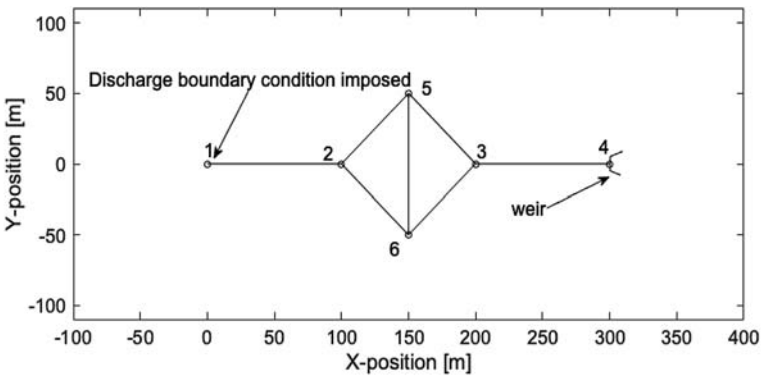

Figure 6.4 Layout of the artificial network. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

length of 100 m (apart from the conduit between nodes 5 and 6, which is $100 \times \sqrt{2}$ m long), with a diameter of 0.5 m. The weir at node 4 has a length of 3 m and a weir level of 8 m above reference level.

At $t = 0$ the system is empty, the boundary condition at node 1 is a time varying discharge (Figure 6.5). For the sake of simplicity, only two model parameters are considered: the hydraulic roughness k_n and the weir coefficient C.

As can be seen in Figure 6.6, the water level at all nodes is more sensitive to a change in the value of the weir coefficient C (continuous line) when compared to the hydraulic roughness k_n (dashed line), i.e. larger absolute values of the elements in the Jacobian matrix. This implies that more information can be obtained related to the value of C when compared to the value of k_n. Node 4 (closest to the weir) shows a high

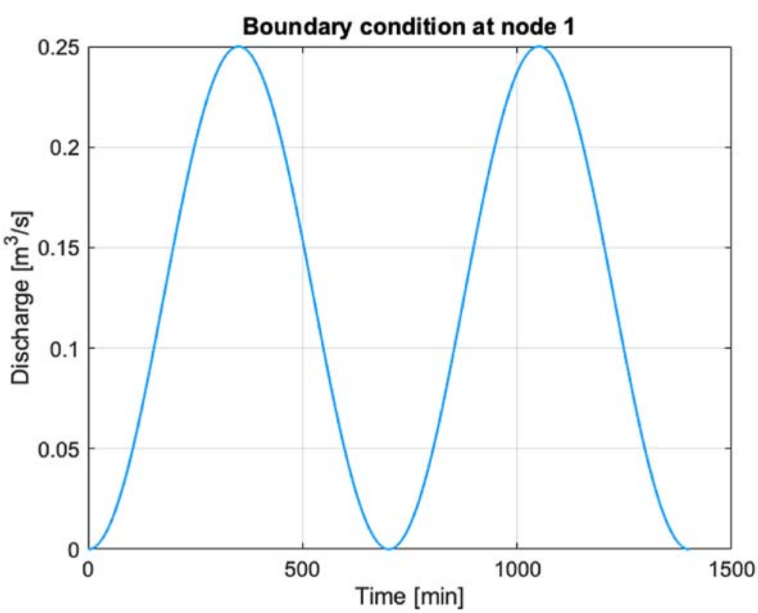

Figure 6.5 Boundary condition at node 1. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

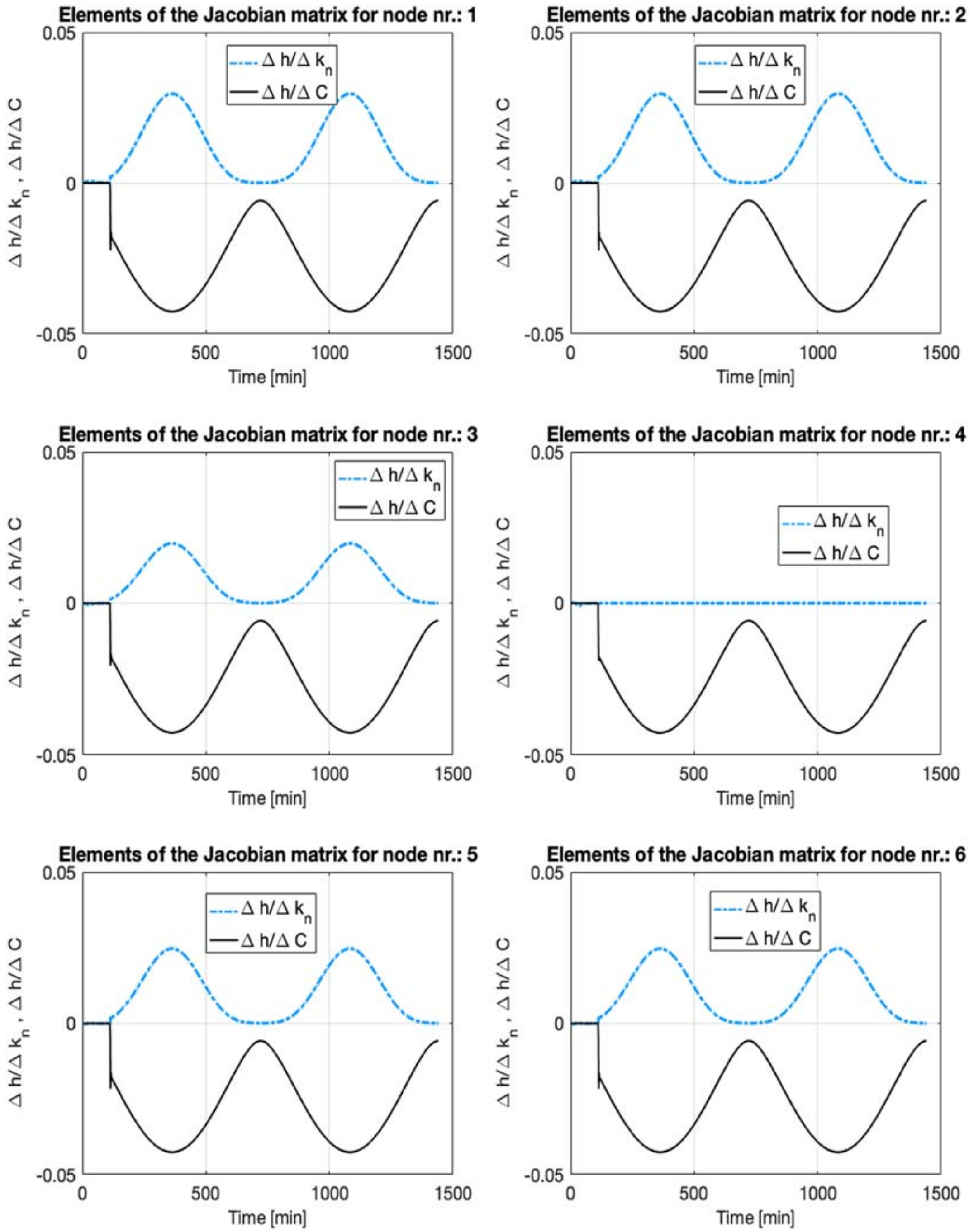

Figure 6.6 Jacobian elements for nodes 1–6. The results for node 2 are enlarged for the time interval [0,160] in Figure 6.7, as some instabilities and obviously wrong results are obtained. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

sensitivity to variation of the value of C and no sensitivity to variation in k_n. This can be understood in the following manner: the water level response at node 4 is almost completely determined by the presence of the weir, while the water level at the other nodes is influenced by both the weir and the hydraulic losses in the conduits. So, node 4 would be an obvious choice for a measuring location as unbiased information for C is obtained (this is only true as long as the boundary condition at node 1 forces the water to flow in the direction of node 4, when flow reverses, the information obtained at node 4 is again a mix between information on C and on k_n. This illustrates the need for using several loads (storm events) when using a model to identify potential measuring locations). With respect to identifying k_n one could pick manhole 1, as this manhole is the farthest away from the weir and the water level variation is influenced by the hydraulic roughness of the conduits over the maximum length. Combined with the unambiguous information for C obtained at node 4, this allows the identification of k_n. Any combination of manholes will result in a certain covariance between the parameter values obtained for C and k_n, as the change in water level is influenced by both parameters. A further observation is that choosing manholes 5 and 6 or just one of them makes no difference in the information gained: the water level in both manholes is identical and responds identically to a variation in C and k_n. This is the same as stating that the water levels at both locations are one-to-one correlated. This implies that on the one hand acquiring information from manholes that are in a hydraulic sense 'neighbours' does not add substantial amounts of information. On the other hand, it allows for consistency checks between the two sensors. An expert will look for manholes that on the one hand provide information but on the other hand, preferably, also show some (limited) overlap in the expected validity range of the parameter which enhances the data validation options. Such an overlap in water levels can easily be checked using a model.

There are some other observations that can be made from the model results:

- The absolute value of the elements in the Jacobian depends strongly on the discharge (Figure 6.6).
- In the time window between $t = 100$ and $t = 120$ minutes, 'artifacts' show up (Figure 6.7): when the water level at node 4 reaches the weir level, a brief time window follows in which some small oscillations in discharge occur resulting in unrealistic values for the Jacobian elements.

These issues are due to the limitations of the software, and some incorrect modelling (in this example this was done on purpose to illustrate the effect): settings of the numerical solver, choice of time steps, etc. Further it has to be realized that in the determination of the Jacobian elements, a finite difference approach is used for quantifying a gradient: when small gradients occur, even round off errors may induce unrealistic values for the gradients.

Overall, a model can be a very valuable aid when choosing monitoring locations, although when doing so the following issues are to be kept in mind:

- Be aware of the occurrence of numerical 'artifacts' (even in very simple networks, as shown previously).
- Make sure to use a range of loads (storm events), as the sensitivity of model results (e.g. water levels) to a variation of model parameters related to any $Q(h)$ relation strongly depends on the load, therefore one is well advised to use a series of loads to judge the potential of all relevant nodes to be a monitoring location.
- Be sure all processes related to the calibration parameters chosen do indeed occur in the model simulation.
- The results of the sensitivity analysis (the resulting Jacobian matrix) can, and in the general case very likely will, vary with varying initial parameters.

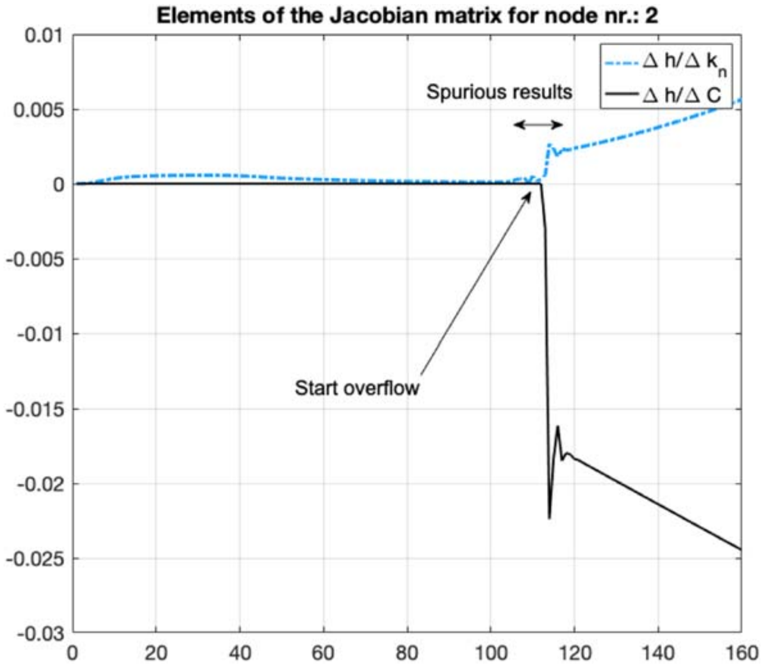

Figure 6.7 Anomalies in the time window between $t = 100$ and $t = 120$ minutes (onset of the overflow). *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

The information obtained by estimating the elements of the Jacobian matrix can be used in different ways to determine the optimal locations of sensors. In addition to a simple approach as explained above, more sophisticated methods could be employed to use this related information. These include statistical theory and information theory-based methods. Examples of statistical theory methods include:

- *Alphabetic design methods* that are based on conventional statistical theory and linearization of the system models in question. Among the most used alphabetic metrics are the D-optimality, A-optimality and V-optimality metrics (e.g. Kapelan *et al.*, 2003, 2005).
- *Bayesian theory-based methods* that seek to maximize the gain in information between the prior and posterior distributions of parameters, inputs, or outputs.

In addition to statistical theory methods, information theory-based methods can be used to locate sensors. These methods work by maximizing the information content that sensors can provide. Examples include:

- *Entropy based method* based on the principle of maximum entropy, i.e. a selection of measurement locations and other sampling variables that result in best current knowledge about the observed phenomena.
- *Mutual information-based methods* that use the amount of information that one variable of interest contains about another variable of interest. An optimal sensor network design should avoid collecting repetitive or redundant information, i.e. it should reduce the mutual (shared) information between sensors in the network.

- Other methods such as the *value of information based method* (based on the value a decision-maker is willing to pay for extra information before making a decision where to locate sensors), *fractal-based method* (that utilizes the concept of Gaussian self-affinity to measure the dimensional deficit between the observations of the analysed system and its real domain) and *network theory based method* (makes use of three variables, average clustering coefficient, average path length, and degree distribution to distinguish between different sensor network designs).

More details about the methods mentioned above can be found in Chacon-Hurtado *et al.* (2017).

6.2.5 Timescales, sampling frequency and measuring uncertainty

Assessing the timescales at which some of the processes show significant changes in time provides a first indication of the order of magnitude of the sampling frequency needed.

Table 6.1 provides a first indication of the timescales of major processes encountered in urban drainage systems (Schilperoort *et al.*, 2012). These values, of course, deviate in specific cases but they can serve as a rule of thumb in the preliminary stages of the design.

In most cases it is not just one process that is of importance to the monitoring goal. In these cases, theoretically, one could decide on a varying sampling frequency (e.g. switching from one sample per hour under dry weather flow conditions to once every 5 minutes when a storm event is detected). This approach was applied in the past (1980s–1990s) when data transmission and storage capacity were expensive and frequently unreliable. The reasons for doing so have become less and less valid as data handling has shown considerable improvement in recent decades. Therefore, one is well advised to pick a constant and uniform sampling frequency throughout the monitoring period for all sensors involved, defined by the process with the smallest expected characteristic timescale.

Table 6.1 Range of characteristic timescales of some relevant hydraulic and hydrological processes in urban drainage and stormwater management.

Process	Timescale
Dry weather flow in foul- and combined sewer systems	1–4 hours. For larger catchments (i.e. >100 000 inhabitants) the timescales tend to be at the higher end of the range, and for smaller catchments towards the lower end.
Storm induced CSO events	1–15 minutes, strongly depending on the system under study and the storm event
Infiltration	Minutes-hours, depending on the soil type, initial conditions, climate
Rain	Minutes-hours, depending on the type of storm event (intense summer convective showers to frontal storm events)
Evaporation	Hours-weeks-months
Flow in storm sewers	Minutes-hours, strongly depending on the forcing storm event
Emptying of system using pumps or outlet	1–10 hours
Run-off	See storm events (minutes-hours)
Clogging of filters (SUDS – sustainable urban drainage systems)	Months-years

There are other advantages of using a constant time step for monitoring. Most methods for analysing time series assume equidistance (fixed time interval) between data points, as time series analysis is an important tool in the assessment of data quality and data validation (see Sections 9.3 and 9.4). This important aspect of monitoring becomes easier and more straightforward to implement when choosing a constant and uniform sampling frequency. It has been shown in practice that implementing rules for switching sampling frequencies adds significantly to a decrease in the reliability of the monitoring system reflected in the data-yield (after data validation, see Chapter 9) of the monitoring networks. For example, the case where the detection of rain triggers a switch in sampling frequency of e.g. a water level sensor results in loss of data for water level measurement when the rain sensor is not correctly functioning. In other words, adding complexity to the monitoring system should be considered only when the consequences with respect to data- and information yield strived for are considered. Another general issue to reckon with, when choosing a sampling frequency, is that reducing the sampling frequency is a decision that can be made on evaluation of historical results from a given monitoring station, increasing the sampling frequency can only be made based on 'gut' feeling. Therefore, it is suggested to start a monitoring project with a sampling frequency as high as practically feasible and decide on reducing that frequency after a first evaluation of the data obtained from the system. For more information on reasons for data/information loss, the reader is referred to literature (e.g. Schilperoort *et al.*, 2008 and Chapters 2, 3, 4, 5, 9 and 11 of this book).

6.2.5.1 Application of a model to quantify timescales of a system

In cases where a model of the system that mimics the process(es) one is interested in is available, such a model can be used as a means for obtaining information for picking a sampling frequency. It has to be realized however that such a model is not (yet) calibrated and may contain serious deviations from the 'real' world due to missing data, wrong data, etc. Further all methods described in the following sections are more or less sensitive in their outcomes to the quality of the model, the loads chosen (storm events) and model parameter settings (see Section 6.2.4). The implicit underlying assumption here is that, in close resemblance to the famous story of Baron von Munchhausen (Raspe, 1895), one can bootstrap him/herself out of the swamp: i.e. in an iterative process the model is improved using monitoring data, using the improved model the monitoring design can be improved, etc. To the authors' knowledge, such a scheme has not been reported in the literature to be applied so far, and no guarantee can be given for the iterative process, as described, to converge.

6.2.5.2 Upper and lower limits of the sampling frequency related to measurement uncertainty

When taking samples from a process and the demand is that the process can be reconstructed from the data, a lower and upper limit for the sampling frequency can be defined in relation to the uncertainty of the sampled parameter values. Loosely defined the lower limit is that sampling frequency below which information on the process is lost, while the upper limit is defined by that sampling frequency beyond which 'noise' is being monitored. Both limits depend on the process characteristic timescale(s) and the uncertainty of the monitoring data. The former is to a large extent a given fact and is known in terms of order of magnitude only, while the latter is largely a matter of choice (design).

6.2.5.2.1 Lower limit

The lower limit of the sampling frequency can be quantified using a time-domain analysis on either a measured time series or a time series obtained from a model. The reasoning is as follows (for details see Lepot *et al.*, 2017).

Suppose an equidistant time series describing some (hydraulic or hydrological) process $x(t)$, a value $x(\tau)$ with $t_{j-1} < \tau < t_j$ is obtained from a simple linear interpolation (Equation (6.2)):

$$x(\tau) = \alpha x(t_{j-1}) + \beta x(t_j) \tag{6.2}$$

where α and β are weighing factors in the interpolation.

Let us assume the process has a known variance σ_x^2 and mean value μ_x, and the normalized autocorrelation function of the process $\rho_x(\tau)$ is known as well. In that case, the mean squared error introduced by the interpolation process is defined by Equation (6.3):

$$mse_{interp} = \frac{1}{2}\sigma_x^2[3 + \rho_x(t_{j-1}, t_i) - 4\rho_x(t_{j-1}, \tau)] \tag{6.3}$$

From Equation (6.3) it is easily checked that for a constant process (no change in time so $\rho_x(t_1, t_2) = 1$ for any combination of t_1 and t_2), the interpolation error is zero. In this case, theoretically, one measurement over the whole monitoring period would suffice, on the other hand, for a completely random process ($\rho_x(t_1, t_2) = 0$ for any combination of t_1 and t_2), the interpolation error (more correctly the introduced uncertainty) $\sigma_{interp} = 1.5^{0.5}\sigma_x \approx 1.22\sigma_x$. The autocorrelation function is defined as the correlation between the time series x and the same series x, but shifted over a certain time shift T (Equation (6.4)):

$$r(T) = r(t_{i-T}, t_i) = \frac{1}{n-1}\sum_{i=1}^{i=n}(x_{t-T} - \mu_x)(x_t - \mu_x) \tag{6.4}$$

The normalized autocorrelation function is defined by Equation (6.5):

$$\rho(T) = \frac{r(T)}{r(0)} \tag{6.5}$$

The demand set on the allowable interpolation error is given in Equation (6.6):

$$mse_{interp} < \gamma\sigma_m^2 \tag{6.6}$$

where σ_m is the standard deviation of the measuring uncertainty of the chosen measuring device, and γ is a multiplication factor: $\gamma = 1$ implies that the uncertainty introduced by interpolation is equal to the uncertainty of the measured data. Given a process (and hence an autocorrelation function), setting a value for γ allows the choice of a combination of sampling interval and measuring accuracy so as to comply with the demand set in Equation (6.6).

The two sources of uncertainty (interpolation and measuring) may be assumed to be independent, which implies that the uncertainty interval (95% coverage interval, see Section 8.2.3) of a value x that is based on interpolation between two data points is the sum of uncertainties due to the interpolation process, and the measuring uncertainty ($\sigma_x^2 = \sigma_m^2(1+\gamma)$) in the data itself is $[x - 1.96\sqrt{1+\gamma}\sigma_m, x + 1.96\sqrt{1+\gamma}\sigma_m]$.

One has to realize that this result has, strictly speaking, no generic validity as it is based on the characteristic of the process that is assumed in the analysis. This implies that in order to achieve a robust result, the analysis should be made for a range of mutually different events, as the dynamics of the event largely define the autocorrelation function.

6.2.5.2.2 Upper limit

Strictly speaking, the upper limit of the sampling frequency is determined by the progress of sensor innovation. However, in practice there is also a maximum sampling frequency above which no additional information is obtained. This frequency depends, likewise for the lower limit, on the process to be monitored and the uncertainty of the sensor readings.

The signal to noise ratio (*SNR*) is a key parameter when defining an upper limit for the sampling frequency. The *SNR* is defined as in Equation (6.7):

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{6.7}$$

where P_{signal} is the power of the signal one wants to monitor and P_{noise} is the power of the measuring noise.

In general literature, an *SNR* of 3 is used to define the limit of detection (LoD) (e.g. Desimoni & Brunetti, 2015). The power of a signal depends on the signal's frequency and amplitude, it is therefore natural to revert to an analysis in the frequency domain. To this end, both signal and measuring noise are Fourier transformed. When the measuring noise is modelled as white noise (mean zero, constant variance $\sigma_n{}^2$, autocorrelation is zero), the power of the signal is constant over the full frequency domain and is equal to the variance of the signal. This implies that by setting a value for the desired *SNR* and a known value for the power of the signal to be detected, the maximum allowable value for the noise level is obtained, which is translated directly back into a demand to be put on the allowable uncertainty level of the monitoring set-up (largely the choice of the sensor) in conjunction with the sampling frequency. When applying a Fourier transform of a signal there is an upper limit to the frequency that is the well-known Nyqvist frequency. The Nyqvist frequency f_N is defined by the fact that at least two sampling points are needed for each frequency to reconstruct the signal. Therefor f_N is given by Equation (6.8):

$$f_N = \frac{1}{2\Delta t} \tag{6.8}$$

where Δt is the sampling time interval.

Adding frequencies above this limit results in 'aliasing'. In most software development environments, e.g. Matlab®, pre-programmed functions that take this limit into account automatically are available. Parts of the signal having frequencies larger than the Nyquist frequency cannot be identified from the measurements.

The underlying principle of a Fourier transform is that a signal $f(x)$ can be written as a summation of periodic functions (Equation (6.9)):

$$F(k) = \sum_{j=1}^{j=N_e-1} f(j) e^{\frac{-i2\pi kj}{N_e}} \tag{6.9}$$

where i is the imaginary unit. The real part of $F(k)$ is the amplitude belonging to a frequency of $2\pi k/N_e$, with N_e the number of elements in the time series. The inverse transformation (from frequency domain to time domain) reads as Equation (6.10):

$$f(n) = \frac{1}{N_e} \sum_{k=0}^{k=N_e-1} F(k) e^{\frac{i2\pi k}{N_e}} \tag{6.10}$$

The 'power' of a certain frequency in the signal is defined as Equation (6.11):

$$P(k) = \sqrt{F(k)\overline{F(k)}} \tag{6.11}$$

The overlined $F(k)$ indicates the complex conjugate. For practical applications, built-in functions in packages like Matlab® can be used (for Matlab® that is the Fast Fourier Transform, `F(t)=fft(f)` and its inverse `f=ifft(F)`). For detailed information on Fourier transforms and analysis, the reader is referred to the literature on the subject, e.g. in Grafakos (2014).

6.2.5.2.3 Examples

In order to illustrate the theory discussed in the previous two sections, two examples are given. The first example is on how to use the results of a hydrodynamic model to define lower and upper limits of the measuring frequency related to the process characteristics and to the uncertainty of the measuring device. The second example shows the results when applied to a real-world problem.

Example 1. Figure 6.8 shows a hydrograph (given as water level) for two locations obtained from a hydrodynamic model for a combined sewer system. The process variability is expressed as standard deviation and is calculated to be 0.64 m and 0.16 m for location 1 and location 2, respectively.

Figure 6.9 shows the autocorrelation functions of the hydrographs (the Matlab® commands used for determining sample frequencies are `meas_freq_exp.m` and `min_max_sample.m`, available for download at https://doi.org/10.2166/9781789060102).

In the model a time step of one minute is chosen. For the sake of illustration, let us suppose we want to set a measuring frequency of once per 2 min, so a reading at $t = 0, 2, 4, 6\ldots$ minutes and find out which maximum uncertainty in the measuring device is acceptable to meet the condition $mse_{interp} < 3 \times \sigma_m{}^2$ (Equation (6.6)). For reconstructing by interpolation the value of the water level that occurs at $t = 1, 3, 5\ldots$ minutes, apply Equation (6.12):

$$\frac{1}{2}\sigma_p^2[3 + \rho(0.2) - 4\rho(0.1)] < 3\sigma_m^2 \tag{6.12}$$

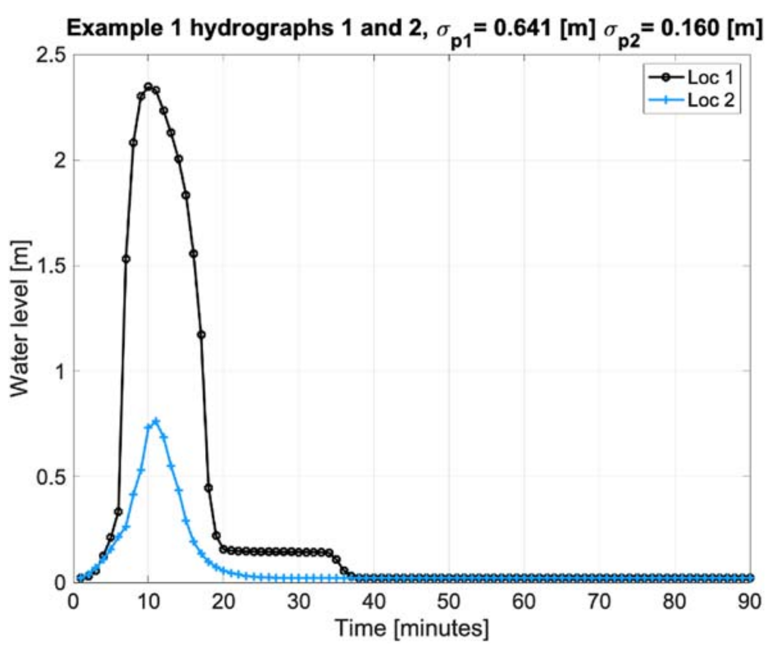

Figure 6.8 Examples of hydrographs. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

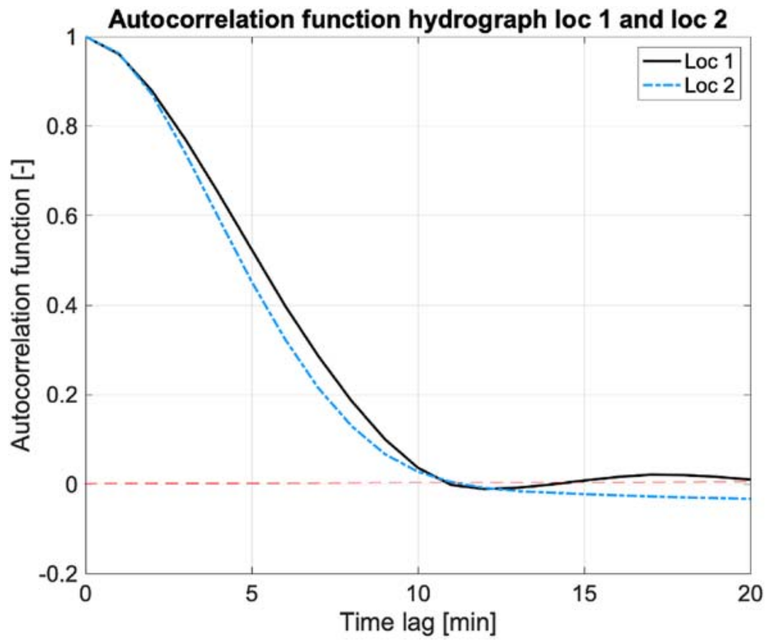

Figure 6.9 Autocorrelation function for the hydrographs given in Figure 6.8. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

Inserting the corresponding values of the autocorrelation functions results in Equations (6.13) and (6.14):

$$\sigma_{m,1} > 0.048\ m \tag{6.13}$$

$$\sigma_{m,2} > 0.036\ m \tag{6.14}$$

So, in this case, for the two locations chosen, no information loss due to interpolation occurs at a sampling rate of once per 2 minutes, provided the uncertainty of the measuring device is less than 0.036 m, which is not a very serious constraint for this case. As the hydrographs, and hence the values of the process variance and the correlation function, are specific for the locations and the load (storm event) imposed on the model, this result can only be regarded as valid for this particular location under the specific load imposed on the model. In order to obtain a more comprehensive indication of the minimum interpolation, an alternative reasoning starts with choosing the measuring uncertainty and works out the maximum allowable time interval between measurements.

In Figure 6.10 the power spectra for both hydrographs are shown. As the time resolution of the hydrographs is one minute, this limits the maximum sampling frequency that can be derived. If one wants to be able to reproduce the signal (the hydrographs), the 'power' of the measuring uncertainty should be less than the smallest value in the spectra of the hydrographs, or more precisely the signal to noise ratio should be such that $SNR > SNR_{min}$. From Figure 6.10, it follows that with $\gamma = 1$ the maximum allowable measuring uncertainty, given the hydrographs and given the frequency of 1/min, is 0.00096 m, which for practical purposes is set at 0.001 m, resulting in a 95% confidence interval of 0.004 m. In an alternative approach, one can chose the measuring uncertainty and determine the 'cut-off'

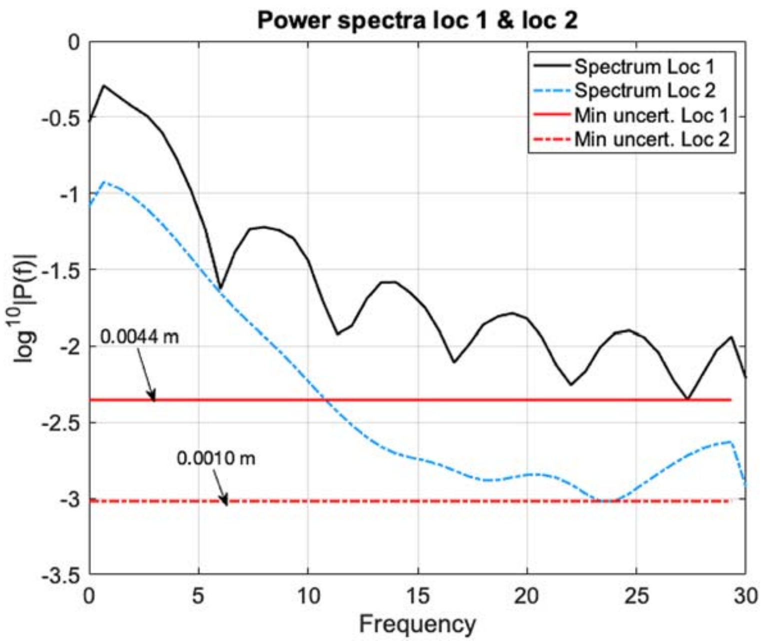

Figure 6.10 Power spectra for the hydrographs given in Figure 6.8. *Source*: Francois Clemens-Meyer (Deltares/TU Delft/NTNU).

frequency (i.e. that frequency of the signal with a power that fulfils the demand $SNR = SNR_{min}$ determining the maximum frequency that corresponds to the chosen SNR_{min}).

Example 2. The example for determining sampling frequencies is based on a study reported by Clemens (2001), for a small catchment with a combined sewer system, known as 'De Hoven' in the centre of the Netherlands. A preliminary rainfall-run-off model and a hydrodynamics model were available. Using the two approaches presented in Section 6.2.5.2 to identify the minimum and maximum sampling frequencies, 10 design storms were used to check for 'load dependencies' of the results. The monitoring set-up was realized and after obtaining data from the monitoring network, the preliminary model was calibrated. A comparison was made between the values obtained in the design phase and what would have been the results when the calibrated model was available during the design (validation using observed storm events). In order to show that the specific process characteristics do not only depend on the storm, but on the location in the system as well, the example is presented for three monitoring locations. The value for γ in Equation (6.6) is set to 1, and the standard deviation of the measuring uncertainty is chosen to be 0.003 mm.

From Tables 6.2 and 6.3, some interesting conclusions are drawn as follows. Per storm, mutually different values for the max and min sampling interval are obtained, but the variation is within a factor of maximum 3. With increasing return period T_r (implying increasing rain intensity, van Luijtelaar & van Rebergen, 1997), the value of the sampling intervals decreases (obviously triggered by an increase in speed of variation of water levels/discharges forced by larger rain intensities). All design storms resulted in a CSO event, while such an event occurred for only two out of the five storms used in the evaluation. It can be seen as well that there is a significant difference in the intervals obtained for real events 1 and 5

Table 6.2 Maximum sampling interval (minimal sampling frequency) for three locations and 15 storm events (10 used for design with T_r the return period, 5 for validation afterwards).

Storm event	Remark	Δt_1 (s)	Δt_2 (s)	Δt_3 (s)
01 $T_r = 0.25$ y	Overflow	767	394	554
02 $T_r = 0.25$ y	Overflow	660	303	391
03 $T_r = 0.5$ y	Overflow	561	472	339
04 $T_r = 0.5$ y	Overflow	547	331	332
05 $T_r = 1$ y	Overflow	419	444	394
06 $T_r = 1$ y	Overflow	548	336	341
07 $T_r = 2$ y	Overflow	365	398	392
08 $T_r = 2$ y	Overflow	446	409	345
09 $T_r = 5$ y	Overflow	327	368	382
10 $T_r = 10$ y	Overflow	324	345	370
Real event 1	Overflow	739	838	849
Real event 2	No overflow	1619	1802	1414
Real event 3	No overflow	2524	2095	2298
Real event 4	No overflow	1490	1422	1628
Real event 5	Overflow	1153	1046	1564

on the one hand and the events 2, 3 and 4 on the other. Clearly a CSO event is a 'fast' process setting the limits for the sampling interval. Using the 10 design storms resulted in choosing a sampling interval of 60 seconds, which turned out to be a 'safe' choice as the values obtained from the recorded events all indicate that a larger value would have been sufficient, it has to be noted here that the output time-resolution of the

Table 6.3 Minimum sampling interval (maximal sampling frequency) for three locations and 15 storm events (10 used for design with T_r the return period, 5 for validation afterwards).

Storm event	Remarks	Δt_1 (s)	Δt_2 (s)	Δt_3 (s)
01 $T_r = 0.25$ y	Overflow	158	135	114
02 $T_r = 0.25$ y	Overflow	171	63	60
03 $T_r = 0.5$ y	Overflow	167	72	75
04 $T_r = 0.5$ y	Overflow	130	60	60
05 $T_r = 1$ y	Overflow	95	64	65
06 $T_r = 1$ y	Overflow	134	60	60
07 $T_r = 2$ y	Overflow	61	89	79
08 $T_r = 2$ y	Overflow	105	74	72
09 $T_r = 5$ y	Overflow	60	60	60
10 $T_r = 10$ y	Overflow	60	60	60
Real event 1	Overflow	199	280	375
Real event 2	No overflow	654	675	686
Real event 3	No overflow	1196	658	1082
Real event 4	No overflow	469	527	505
Real event 5	Overflow	329	300	146

software used (Infoworks™) was set at 60 seconds, this choice automatically determines the lower limit of the sampling interval.

6.2.6 Networks of rain gauges

In virtually every monitoring project related to urban drainage and stormwater management, data on precipitation are indispensable. As discussed in Chapter 2, radar based or microwave attenuation measurements (e.g. Fencl *et al.*, 2015) are employed successfully for covering relatively large areas, also in the urban environment, with up to 5 minutes time resolution and, depending on the technology applied, a spatial resolution of $\sim$1 km^2 or less. There is however still a need for validation with ground-based point measurements. The data for the latter is usually obtained from a network of rain gauges. Installing 'the optimal rain gauge network' in an urban environment is challenging for a number of reasons:

- Demands set on the environmental conditions for a rain gauge set-up as formulated in standards (e.g. WMO, 2018) and Chapter 2 of this book), are hard, if not impossible, to meet.
- Considerations on avoiding vandalism dramatically limits the number of usable locations for installing a ground-based rain gauge.
- Microclimate (especially with respect to wind influences) may cause systematic, difficult to access and/or to compensate for, deviations in the readings.
- Accessibility of locations should be guaranteed, while at the same time discouraging access by trespassers.

As shown by Schellart *et al.* (2012), even on very small time and space scales, a significant variability in observed rain intensities occurs which is due to the fractal characteristics of the rainfall process (Sivakumar, 2010). Variability of rain intensity in time and space reportedly affects the results of urban run-off modelling (Bruni *et al.*, 2015; Ochoa-Rodriguez *et al.*, 2015). The high imperviousness and the small-scale heterogeneity of the urban environment result in very short characteristic response times requiring rain data on a wide range of time and spatial scales. In many cases, though, the demands put on the rain data are not that strict. For example, when discriminating between 'dry' and 'wet' days, e.g. for applying a simple method for determining infiltration rates in drainage networks on a catchment scale (see Section 3.5), one could argue that when an estimate of the daily rain volume is obtained, this would suffice. For a homogeneous catchment (in terms of elevation, land use, etc.) at least one rain gauge in the centre of the catchment would be enough, provided the catchment is small ($\sim$1 km^2). For redundancy reasons, it is suggested to install at least two rain gauges in two close, though separate, locations (e.g. Chartered Institution of Water and Environmental Management [CIWEM], 2017). This allows for cross validation when analysing the data. When a long-lasting systematic deviation between the two observations persists, this may hint at either an issue with the sensors (not uncommon), or indeed a significant difference in the micro-climates between the two locations.

When setting up a monitoring network for rain intensity for a larger urban area (typically larger than 1 km^2), the macro design is more complicated. A first step would be to subdivide the area under study into regions with similar characteristics in terms of land use, presence of vegetation, etc. Further, when the area is hilly, differences in rain microclimate can be expected when there is a predominant wind direction that induces rainfall at the windward side of a hill and a relatively dry microclimate at the leeside. Having done that, a (sub)network per region has to be designed. Basically, this is the same problem as picking locations for water level measurements or discharge as discussed in the previous paragraphs. A reasoning along the same lines is therefore at hand:

- Eliminate all locations that are not eligible for a range of practical reasons.
- Depending on the available budget, decide on the feasible maximum number of gauges.
- Distribute the *N* possible sensors over the *M* potential locations in such a manner that the variability in a spatial sense is accounted for, given the measuring frequency.

It is obvious that the results of the two first actions mentioned depend to a significant extent on the choice of the measuring system and the associated costs for maintenance and operation. The reader is referred to Chapters 2, 7 and 8, respectively, for details on measuring principles, maintenance, and calibration and uncertainty assessment; aspects that affect the choice of the measuring equipment and implicitly their demands in terms of services and environmental requirements.

Grounds for eliminating potential locations for rain gauges are summarized as:

- The presence of leaf abscessing trees, as these will cause significant risk of malfunctioning rain gauges.
- A poor accessibility. The site should preferably allow for a safe, though limited, access, i.e. an employee should be able to safely enter the site. At the same time however, trespassing should be discouraged as much as is feasible to avoid vandalism or theft.
- The non-availability of basic services (i.e. power supply, data communication).
- The presence of tall buildings that significantly influence the local wind field (strength and direction).

Regarding methods for obtaining an 'optimal' network, given the potential locations and a number of locations limited by the available budget, several authors have reported their findings (e.g. Fu *et al.*, 2016; Lei & Schilling, 1993; Rodríguez-Iturbe & Mejía; Shaghaghian & Abedini, 2013; Tiwari *et al.*, 2020). These methods are very diverse in nature and most of them require a substantial background knowledge on mathematical optimization methods, which is beyond the scope of this book to discuss in any detail. The interested reader is referred to the cited references. Ochoa-Rodriguez *et al.* (2019) provide an up-to-date review on both radar rain data and rainfall data obtained from rain gauge networks, and on the influence the configuration/density of the rain gauge network has on the results obtained. Notwithstanding the ongoing developments in the understanding of the properties of rainfall and the manners of recording them, some generic practical insights with respect to network (macro) design are derived from the literature:

- Install at least one, preferably two, gauges per homogeneous (sub)catchment.
- Install at least one gauge per km^2.
- Rain gauges at the periphery of the area under study are a very important source of information on rainfall spatial variability and heterogeneity.
- It is suggested to favour a higher number of measuring locations over a lesser number of locations with high accuracy (and as such more expensive) apparatus when budget issues force a limit on the number of observation stations.

The resolution in time for rain gauges is normally chosen between one day (daily rainfall volume) and one to five minutes, the latter frequency is regarded in practice to be the upper meaningful limit (e.g. for hydrodynamical models: in such a case the level of geometrical detail of the model itself has to be in line with the time-space resolution of the rainfall data), although for research purposes even smaller timescales can be of interest, provided this is combined with a high spatial resolution as well.

When designing a network of rain gauges, a good starting point is to identify properties that provide, or are likely to provide, favourable conditions. Such properties can be:

- Properties owned by the municipality and/or the local water utility, as in most cases these organisations will be involved with the project. Usually, these sites are to a certain extent protected against trespassing and offer a range of basic services for operation.
- Private properties of people involved with the project or other employees from associated parties (consultant, university, municipality, waterboard or involved contractors/construction firms).
- Owned by people having an amateur weather station available. Enquire about their existence and ask them if you can share data. Of course, one cannot expect such people to have top-notch equipment that works with the same protocols as used in the project. Nevertheless, amateur data can be of great value. In some cases, it can even be worthwhile to buy professional equipment and make it available for interested amateurs in exchange for free maintenance and data sharing. Especially in long-term monitoring this can be efficient and effective.

The demands put on the micro design of a rain gauge set-up are discussed in Section 2.2.

6.3 MICRO DESIGN: FROM THE MACRO SAMPLING DESIGN PLAN TO UP AND RUNNING MONITORING STATIONS

The previous section (Section 6.2) provides methods and examples to identify the number of monitoring stations, their locations, and data acquisition frequencies. This section aims to deliver key points to build each monitoring station of the previously designed network, from the sensor selection to the final construction, which must be calibrated and tested to ensure the system is complying with the specifications. After the definition of the goals (Section 6.3.1), Section 6.3.2 suggests a nine-step method for the micro design. Pros and cons of sensor technologies introduced in Chapters 2 to 4 are then detailed with respect to the micro design. Once the monitoring station has been built, Sections 6.3.4 and 6.3.5 introduce few basic tests to ensure the system is running properly prior to its start, and during the first runs, respectively. A few case studies on micro design are briefly presented at the end of this chapter.

6.3.1 Definition of the goals: long-term, mid and short-term installation – 24/7 and event sampling

Once the number of monitoring stations and their locations have been established (Section 6.2), the specifications of each station must be defined prior to the design. To ensure that stations will fulfil the expectations, the goals of each of them should be clearly stated. Several key points deserve some reflection to set reliable specifications.

The first one is the life expectancy of the set-up: a few weeks, months, years up to one or two decade(s). This duration will affect the choice of data communication standards, quality of the equipment and therefore, often, the budget needed, and the time required for the design. For a long-life duration, plausible future needs should be anticipated, data communication standards should still be available until the last refreshment of hardware and software, and a better hardware quality is highly advised to deal with aging. When choosing sensors, one is well advised to balance price, stability and accuracy of the sensor against maintenance, replacements and personnel costs. Ideally one should calculate the cost per data-point of a given minimum quality to make a decision. A temporary monitoring station can be built with lower quality hardware (but not as a necessity) and without consideration of what the future communication standards might be in the coming years. Generally, the design of temporary stations requires less time since reflection on potential refreshment of the station is, in most cases, meaningless.

Table 6.4 Key points for the design of a monitoring station.

		Types of monitoring station	
		24/7	**Event sampling**
Life expectancy	Short term	• Robust components • Stability of the power supply	• Detection of events • Robustness regarding the start/stop procedures
	Long term	• Same as above + quality of the components • Scalability of hardware and software	• Same as above + quality of the components

A second point deserves investigation, regardless of whether the station is designed for the short or long term: 24/7 vs. event monitoring station. A measuring set-up, which does not continuously run, should be designed and built in such a way that the entire system can deal with on and off switches, dry and wet weather modes of use, and automatic and robust procedures to start and stop the entire system. Event monitoring stations are more than likely subject to failures, due to e.g. no suction by a pump when the system starts, pump failures (if the system starts and stops too often due to too sensitive a procedure), biofilms on sensor cells, aging and heat of sensors (Table 6.4). Consequently, such stations should be avoided or, at least, be very carefully designed.

6.3.2 Definition of the needs: hardware, software, maintenance, trained people

Based on the goals defined and the key points to consider, the needs must be clearly identified to ensure the budget and the staff are consistent with the goals. If a main rule for a monitoring station could be stated only once, the authors suggest this one: 'The cheapest investment is not always, in the end, the cheapest solution.'

This remark, which is rather explicit, is to be kept in mind when making decisions regarding hardware, software and maintenance. If, at first sight, cheap components seem to be cost effective solutions to build and run a station given the budget allowance or constraint, it is not always that obvious. Cheap components often suffer from instabilities or frequent failure, either increasing staff costs (to solve the problem) or reducing the data yield of the monitoring set-up. Cheap sensors are often built with cheap components not really able to deal with aging under harsh conditions (e.g. humid atmosphere or temperature variations). This reflection is based on feedback received and personal experience of the authors. The risks mentioned are difficult to assess, as there is a lack of data regarding failure statistics of hardware and the cost of maintenance/reparation is too dependent on the staff/material/travel costs to derive a general rule. Given the development of technologies, complete tests and feedback will most likely be available soon in the literature.

The procedure presented in Figure 6.11 can be recommended for the entire design process. Details for each step are given in the following subsections. The design is organized in nine steps, from the sensors selection to detailed drawing, each step allowing a feedback loop towards the previous ones. For each step, the general idea and a few common sense tips are given, to forecast and avoid potential consequences of incorrect design, poor construction, failures of the components or crisis scenarios.

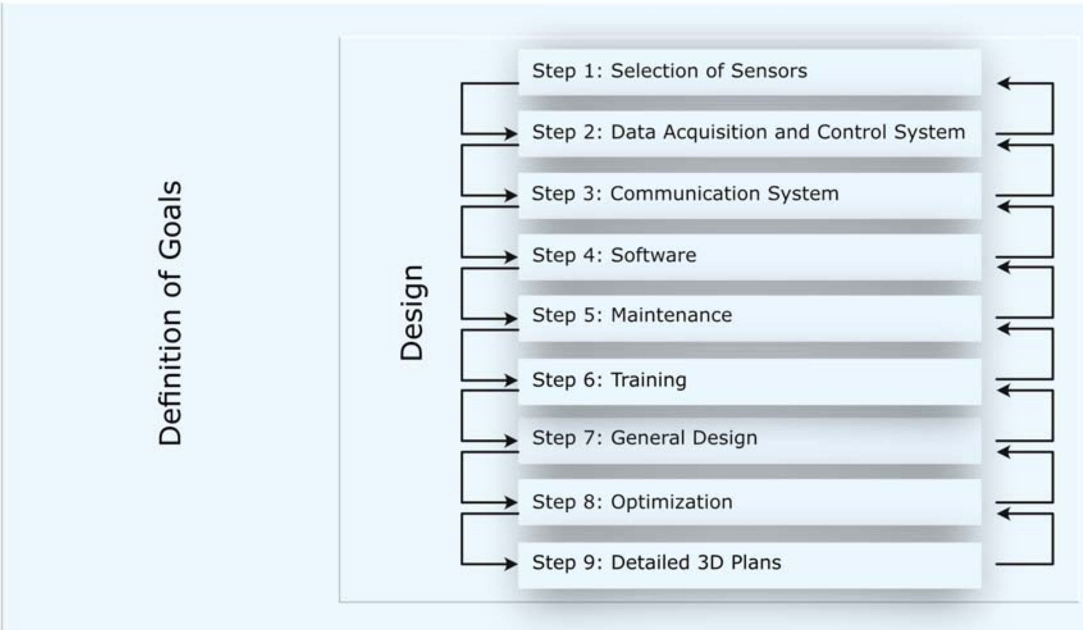

Figure 6.11 Sketch of the recommended nine design steps. *Source*: Mathieu Lepot (TU Delft/Un poids une mesure).

6.3.2.1 Step 1: selection of sensors

For each type of measurement (see Chapters 2, 3 and 4), once the technology(ies) is(are) selected and given the measuring ranges of available sensors, sensors must be selected according to their present and future availability on the market. Given the choice of technologies and manufacturers, the models of each sensor must be selected in accordance with the expected conditions at the measurement location in order to ensure: (i) an overlap between measuring ranges and (ii) a combination of technologies suitable for the expected conditions at the measurement location. If different sensors or technologies are chosen to monitor the same data, but for different parts of the identified range, the total (or even better optimal) measuring ranges should overlap in order to avoid gaps in the complete measurable range by the set-up and increase the data quality. The pros and cons of each type of technology for those sensors are discussed in Chapter 3. The robustness of the probes and the data communication protocols must be considered. The same procedures (except for the measuring range) apply for actuators.

Prior to the choice of the hardware, especially for long-term stations, it is useful to collect information about the system to be monitored: ranges of rain intensity, water level, flow velocity, discharges must be known to obtain information for the sensor selection. Previous or historical data, knowledge, expertise and experience, and modelling must be collected from all reliable sources: system designer, manager, operator, insurance companies, media, citizens, etc. After the verification of those data, the sensors should be chosen according to the expected usual, likely and unlikely conditions that may happen at the measuring location.

Table 6.5 Advantages and disadvantages of different scenarios for sensor redundancy.

Technology	Advantages	Disadvantages
Redundancy with the same hardware	• Comparison between recorded data is easier. • Easier stock management for replacement. • Less skills required to maintain the system.	• The selected sensors should fit the needs, in terms of measuring ranges and conditions. • Sensitivity to unexpected conditions.
Redundancy with the same technology	• Wider measuring range but more sensitive to variations in the measuring conditions.	• Sensitivity to a few special measuring conditions.
Redundancy with different technologies	• Wider measuring range and more robust system regarding the measuring conditions.	• More skills required (for each technology). • More difficulty in data validation.

For data validation purposes and to reduce the likelihood of gaps in data, redundancy (i.e. double or triple sensors for the same data) is one of the key points. Three redundancy solutions can be adopted: (i) by using exactly the same hardware (technology, brand, model), (ii) by using the same technology but different sensors (different brands and/or models), or (iii) by using different technologies. Each of them has advantages and disadvantages, as summarized in Table 6.5.

Ideas for sensor selection

- I 6.1: *Redundancy* – Double or triple the sensors for the same measurement.
- I 6.2: *Variety* – If possible, select different measuring principles for the same measurement, for reducing sensitivity to unusual or unexpected conditions or reasons for failure.
- I 6.3: *Site knowledge* – Grab data and feedback about the site to know the usual, likely and unlikely expected hydraulic and operating conditions.
- I 6.4: *Overlap* – Select your sensors to ensure their measuring ranges are overlapping (e.g. two water levels with the following measuring ranges (0.01 to 0.5 m) and (0.4 to 1 m) – this choice allows measurements between 0.01 and 1 m, with an overlap between 0.4 and 0.5 m).

6.3.2.2 Step 2: data acquisition and control system

Based on the selected sensors (especially their transmitters), the components of the data acquisition system must be chosen. Transmitter outputs (analogue, digital) give strong constraints on the type of data acquisition card(s) or data logger(s) to record the transmitted data, and therefore, on the need for a computer. If required by the design, the same approach remains valid for the choice of control card(s) for actuators. During the selection of the data acquisition system components, an update of the selected sensor(s) might be required to optimize the design (space, cost, availability of hardware): a feedback loop toward step 1 is possible.

Ideas for acquisition and control hardware selection

- I 6.5: *Back to standards* – Be conservative but non-nostalgic and avoid the brand-new promising plug or protocol from a manufacturer not yet tested and evaluated.
- I 6.6: *From now to future* – Try to envisage what standards would be available at the time of the last expected update of the station (this can be done by e.g. keeping track of developments in legislation and/or standards).
- I 6.7: *Scalability* – Select hardware that will allow you to add potential sensors or/and controllers.
- I 6.8: *Modularity* – Select hardware and protocols that will allow you to change individual sensors or/and controllers.

Given the life expectancy, especially for long-term stations, the hardware must be updatable until the last planned refreshment, i.e. the used communication protocols and plugs should be available until this day. For example, lately, the RS232 or RS-485 are almost no longer in use. At the same time, the ethernet ports did not change. The situation is not that clear for USB (Universal Serial Bus): the standards A and B did not change with the generations, the mini-USB (mini A, AB but not B) have been replaced by the micro-USB and, since 2014, the USB-c seems to have become the new standard plug.

Careful attention must be considered for all the components: non-forecasted planned obsolescence might create problems, especially since there is barely any second-hand market for sensors. Technology watchfulness and discussions with experts are the two main means of ensuring the correct decisions regarding this issue. Furthermore, even though it is common sense, proprietary or specific connectors must be avoided.

A communication network is not mandatory if the data can be stored *in situ* and regularly collected. However, the selection of hardware depends on the selected communication network: GSM, internet, radio, etc.

6.3.2.3 Step 3: communication network

The choice of communication network strongly depends on the service availability (Chapter 7) and the selected hardware for the acquisition, and if needed, the control of the station. Maybe the internet (if available) still remains the easiest solution and the most widely applied. The cost and availability of GSM communication is too dependent on the country, the specific location and the supplier to provide a general recommendation. GSM communication might be the best if there is a need for alarm generation and if the staff in charge of the maintenance do not have smartphones. The emerging LoRaWAN (long range wide area network) protocol has been applied successfully in urban hydrology (Blumensaat *et al.*, 2018, 2019; Ebi *et al.*, 2019; Orfeo *et al.*, 2018).

Ideas for communication network

- I 6.9: *Availability* – Which networks are available at the location?
- I 6.10: *Cost* – Actual and forecastable costs.
- I 6.11: *Quantity* – Amount of data to be transferred.
- I 6.12: *Energy* – Energy consumption.
- I 6.13: *Redundancy* – Does the system require redundancy in the communication network?

6.3.2.4 Step 4: software and data storage solutions

The choice of software strongly depends on the selected hardware and the life expectancy. For a short term and simple monitoring station, there is much more flexibility. Usually, such a set-up will not be updated: software updates and availability do not play a relevant role on the choice. Furthermore, if the system will be designed, built, and run by the same people, and if the station is equipped with sensors from the same brand, the choice is only dependent on the skills of the people in charge of the station or the software provided by the sensor manufacturer. However, in the last case, proprietary software often has two serious drawbacks: the lack of information on internal routines and a low flexibility to develop additional routines.

For long term, multi-user and/or multi-sensor manufacturer set-up, the choice of the data acquisition software becomes more critical. The following questions require some deep thought prior to the final choice. Is there a software which can deal with all the sensors and data communication protocols? Is there a software for which the skills and knowledge can be kept over the entire life expectancy, despite turnover of staff? Is there a software which will be updated and available over the same period? If one or more software(s) is(are) an output of the last three questions, the selection is easy. If not, the previous answers should be weighted and analysed to make the final selection. This choice will lead to a widely applied software: there will be people available on the job market with the desired skills; the software will most likely be regularly updated and will follow the change of sensor communication standards.

However, the choice of a new but not widely used software is an option. In such a case, special attention should be devoted to keeping the skills within the organization responsible for the system operation, thus requiring careful management to ensure employees with those skills will stay and/or appropriate training of other staff members is planned.

As for the previous steps, a feedback loop might be activated if the compatibility with the selected hardware is not complete.

Data storage and expected uses of data must be carefully studied during the design step. The following questions should be answered: who will use the data (internal and/or external uses)? What are the skills and knowledge of all the potential data users? Should the data be accessible 24/7 or not? Where must the data be stored and backed up (on site, at the office, in the cloud)? A few common sense tips must be clearly stated.

The first one is associated with the metadata, especially if a network of monitoring stations is set up. Metadata are mandatory to understand the data (importance of metadata is discussed in Chapter 10) and may include at least the following ones: name of the monitoring station, GPS coordinates, type of time used (local or GMT – highly recommended to avoid problems during winter and summer time shifts or with e.g. day-saving time shifts when synching time series from mutual different sources), time step, hardware list (brand, model, serial number), life sheet of all sensors (see Chapter 7) with maintenance data such as: date and type of intervention (including calibration, verification, cleaning, uncertainty data, failures and repairs, after sale service, replacement, etc.), name (action and technician) and comments on each maintenance action, date and calibration/verification data (starting date and validity period). This list can be completed with other topics: their relevance should be discussed with all the people involved in the project.

The second one concerns the data storage location and accessibility (see Chapter 5). On site storage is the easiest method to implement but presents a few drawbacks: such data can be lost and/or stolen on site, are not easily accessible, and require more staff costs for the transfer from the site to the office. A combination of on-site storage and automatic backup at the office is highly recommended. Several options are available to achieve this purpose: continuous or sequential backup (once per hour, per day) while using communication networks (GSM, internet or radio depending on the amount of data to transfer and the corresponding costs).

The third and last one deals with the ease of the data use. Depending on who will use the data and all the plausible mistakes that can be made (changes in the raw data, data deleted, lost, modified, corrupted, etc.), a robust and reliable data management system to ensure the integrity of the data, its backup and a 'user friendly' interface is required. A basic rule: never work with original raw and validated data sets, only on copies.

Software

- CL 6.1: *Cost* – Price of the software and its update?
- CL 6.2: *Open-source* – Or not?
- CL 6.3: *Knowledge* – Available within the team? Widely used?
- CL 6.4: *Compatibility* – With other software or hardware.

Data storage

- CL 6.5: *Separate* – Separate the location of the storage?
- CL 6.6: *Redundancy* – Duplicate the storage hardware.
- CL 6.7: *Read only* – Keep a backup of read-only raw data.
- CL 6.8: *Accessibility* – Easily accessible and efficiently protected.
- CL 6.9: *Experts* – If needed, contact IT experts.

6.3.2.5 Step 5: maintenance

'Plug and play' is most likely the biggest source of frustration in monitoring set-ups. By acting on this commercial promise, a lot of problems can be encountered, especially in urban hydrology, given the constraints related to urban environment and to the (waste)water matrix. In summary, maintenance needs to be considered from the design phase, in order to avoid building an up-to-standard set-up delivering, in the best case, strongly biased data. If monitoring in urban hydrology requires a lot of initial investments, they should not be wasted by insufficient maintenance and data processing costs, which should be evaluated exhaustively. Operation costs to ensure data availability and quality are usually significantly higher than investment costs, especially for long-term monitoring.

The frequency and the duration of maintenance of a monitoring station strongly depend on the type of set-up: number and diversity of sensors, and required maintenance according to the manufacturer specifications (e.g. sensors, pumps, auto-sampler). Combined with the number of monitoring stations of the network, the frequency and duration of maintenance allow estimation of the required maintenance time to be spent *in situ*. A special issue must be pinpointed here: it is not unusual to identify a gap between the required maintenance actions and frequency (claimed or promised by the manufacturer/retailer) and the real ones. It should be no surprise that the claimed required maintenance is only rarely overestimated. On top of this duration, travelling times between the office and the different locations lead to the calculation of the entire required duration to keep the network in good operational condition. In addition to staff costs, trip (fuel, car, insurance, etc.) and consumable (e.g. pump, pipe, calibration standards) costs should be carefully estimated.

Last but not least, the management of the hardware stock needs to be considered. Destruction or malfunction of hardware are unaffordable, and a distinction between critical, sensitive and non-sensitive hardware should be made. On the one hand, pumps and pipes for a by-pass monitoring station can be considered as critical. Data acquisition and control card or hardware (e.g. computer) can be considered as critical or sensitive depending on the goals and the obligations of the set-up. On the other hand, other hardware might be, sometimes, considered as non-sensitive: redundant flowmeters or rain gauges, and calibration standards. The availability and the quality of the data might not be affected if one of these hardware components is temporarily out of order. The key point is data recording and backup. Given the available budget, hardware costs and storage location, a stock of those devices has to be prepared to face upcoming problems and to solve them as fast as the system (and its goals) requires it. A failure risk analysis is highly recommended, e.g. a FMECA – failure mode effects and criticality analysis (Mikulak *et al.*, 2008). Proper management, including an up-to-date list of available hardware and tools, a forecast of the potential needs (e.g. replacement of gaskets every 6 months) and the expected order and delivery times should be planned in advance.

After hardware and software are decided upon, the design of the implementation of all the components must be such that it minimizes the duration of maintenance tasks, i.e. the item to repair or replace should be easily accessible, easy to remove while avoiding (when possible) any disruption in the monitoring. Furthermore, any system that can identify when maintenance operations occur should be designed and set up (e.g. switcher, magnetic contact): such a solution facilitates data validation, while highlighting when the data may be disturbed by the maintenance (e.g. sensors cleaning, calibration).

Additional information on maintenance is provided in Chapter 7.

Objectives to be achieved during the design, for maintenance operations

- CL 6.10: *Operation duration* – Reduce frequency and duration.
- CL 6.11: *Risk* – Minimize the risk for the operators.
- CL 6.12: *Remote* – Allow remote diagnostics to increase maintenance efficiency.

6.3.2.6 Step 6: trained people and training

Since the 'Plug and play' approach is far too optimistic, experienced and trained staff members are essential to ensure the quality of the delivered data. Recruitment and training costs and durations should not be underestimated, especially for complex and long-term set-ups.

After the system is designed, a complete analysis of both existing and required skills and knowledge should be conducted to identify the potential missing ones. Initial training should be planned according to those gaps, while keeping in mind some redundancy between the staff to allow for possible turn overs, like holidays, and ensure there will always be someone able to do the maintenance. For long-term monitoring stations, continuous training of new employees must be set up.

Action to be taken with respect to staff and training

- CL 6.13: *Balance* – Comparison between existing and required skills and knowledge.
- CL 6.14: *Identification* – Identification of suitable training.
- CL 6.15: *Spread* – Spread the skills and knowledge within the team to ensure continuity.
- CL 6.16: *Report* – Report procedures, questions, and comments.

6.3.2.7 Step 7: general design and drawing

In addition to all the foreseeable disfunctions of a system, other unlikely events should be considered to ensure the integrity of the set-up: flooding, thunderstorms and their power supply cuts, vandalism, data communication shutdown, etc. Those are crisis scenarios which require solutions to avoid them or at least minimize risks, damage, and consequences. Table 6.6 summarizes the main solutions for some types of events.

Given the identified possible crises and their solutions, the general set-up and its installation should be designed accordingly. At some locations, regulations and available space might be constraints to this general drawing/layout of the monitoring station: careful attention should be devoted to this step.

Table 6.6 Solutions to avoid crisis.

Type of crises	Solution(s)
Flooding	• Select the right location for the sensitive hardware • Flooding proof set-up
Thunderstorms	• Protection against shutdown • Automatic restart procedures
Vandalism	• Padlock • Alarms • CCTV cameras for dissuasion • Set-up either visible or properly hidden from the surroundings (it is hard to underestimate the creativity and dedication of urban vandals) • Information panel is usually useful to reduce vandalism

Best friends and best enemies

Low- and high-power supplies cabling

A problem that is commonly observed persists: the proximity between low- and high-power cables, i.e. often data transmission and power supply. The alternative high voltage and current of the power supply (typically 220 to 380 V) creates an electromagnetic field. This may affect all the analogue signals delivered by some sensors of the station and, therefore, create some noise in the data. This problem can be encountered with other sources, especially cell phone connections. The proper use of shielded cables (i.e. when the external grid of those cables is connected to the ground) is the most standard and the most efficient solution for this problem.

Electricity and water

Even if it sounds like basic common sense, electricity cables and water pipes should be carefully placed relative to each other. As far as possible, the water should never be in contact with electricity cables: those should be always placed above the water hydraulic head, even in the case of flooding or disruption. The choice of the certified connections (waterproof) helps a lot in reducing plausible consequences of a station malfunctioning.

6.3.2.8 Step 8: optimization

The general drawing is now ready. Can it be optimized? This predesign should be presented to all the people involved in the design, construction, use and maintenance of the monitoring station. Feedback, suggestions, and corrections should be collected, processed, and implemented. In case of hardware failures, especially if the station should run 24/7, the components must be easily replaceable, i.e. the required duration for such an operation must be minimized (see Chapter 7), even if there is(are) (a) redundant component(s). If the acquired data are crucial (for any kind of reason), special attention must be devoted during the design to lead to a fast-to-repair set-up.

Given the general drawing established at step 7, the accessibility of the different components must be carefully studied to minimize the maintenance duration (including safety protocol, e.g. the method to replace components must minimize the need for access within the sewer – always more time consuming and expensive than maintenance on the ground).

Ideas on optimization

- I 6.14: *Collective* – The optimization should be a collective work.
- I 6.15: *User-friendly* – Adopt user-friendly approaches, for any potential type of users.
- I 6.16: *Cost* – Reduce the cost.
- I 6.17: *Question* – Question all the decisions made since the beginning, were they worth it.

6.3.2.9 Step 9: detailed 3D drawing of the monitoring station

After step 8, everything should be clear on paper. The best way to ensure the design reliability is probably to draw in 3D the entire set-up in its environment. This is especially advised when a complex system of cables and pipes is present. A 3D drawing is a powerful means to reveal construction conflicts. This will help the validation of the entire design, can help the authorization that may be required prior to the construction, and may be used for communication purposes with the different actors involved in the project.

3D drawing

- I 6.18: *Dynamic* – Implement and study the dynamics of every component.
- I 6.19: *Sub-contracting* – Those sketches can be used by sub-contractors.
- I 6.20: *Cost* – Reduce the cost.

There are plenty of 3D drawing software packages available, often for free and with comprehensive libraries of objects. The kinematics and the available space for maintenance can be accurately verified at this step (Figure 6.12).

6.3.3 First tests

6.3.3.1 Calibrate

Prior to the data collection, all the sensors and the data acquisition system should be properly checked and calibrated. Chapter 7 is, amongst other issues, devoted to calibration methods and, therefore, this section is

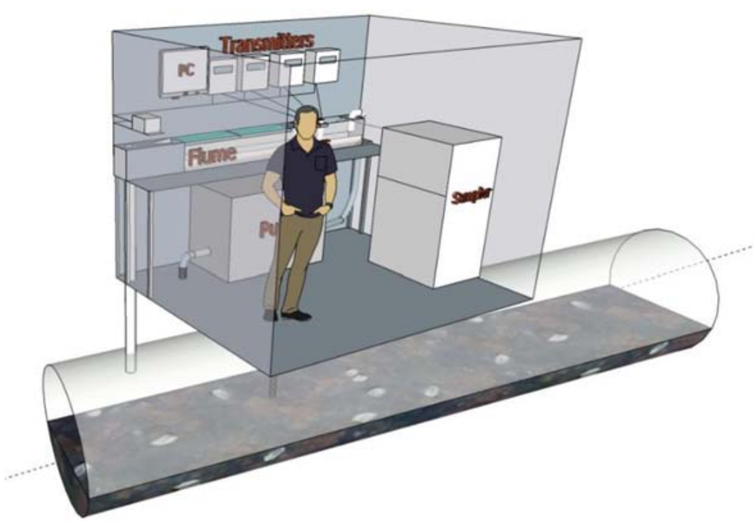

Figure 6.12 3D sketch of the OTHU monitoring station. *Source*: Nicolas Walcker (INSA Lyon).

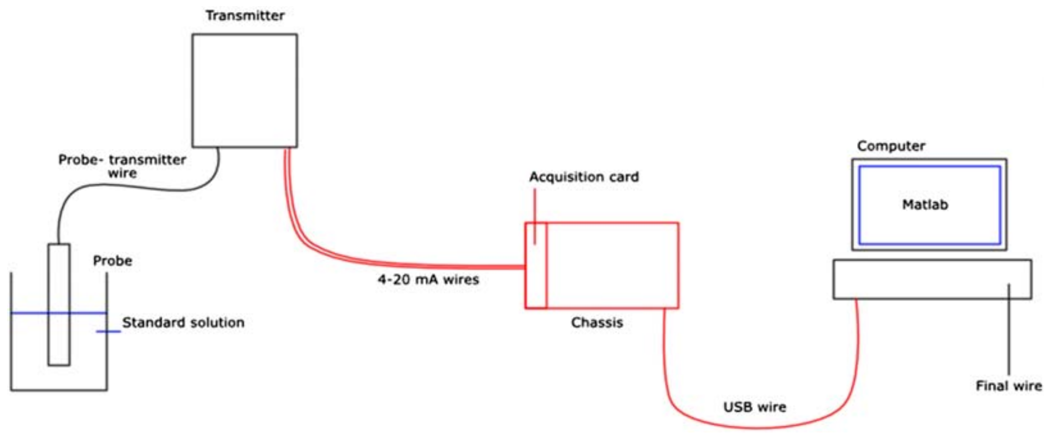

Figure 6.13 Sketch of a measuring chain during an *in situ* calibration process. *Source*: adapted from Lepot (2012).

just a reminder. The calibration of sensors is highly recommended, as a standard best practice, like the periodic checks and re-calibration when necessary (after repair, drifts, etc.).

However, even if sensors are calibrated, the entire data acquisition system should be calibrated as well. In fact, sensors are often calibrated between the measured sample, the sensor and the transmitter (value plotted or recorded within the sensors). Other bias may occur along the measuring chain shown in Figure 6.13.

A typical measuring chain (Figure 6.13) includes different components: sensor, transmitter, data acquisition and control hardware, computer with software and cables. The most frequent calibration procedures are done between the sensors and the transmitter, while using relatively blind internal procedures given by the sensor manufacturer. Bias may also occur between the output of the transmitter and the computer. Therefore, calibration of the full measuring chain is required, from the reference value until the final file (see Chapter 7), not only between the reference value and the screen of the transmitter. In the USA, the standard way to collect data with analogue outputs is the voltage. Between the output of the transmitter and the analogue to digital converter, there are cables. Due to the linear electrical resistance of the cable, the voltage might decrease along the length of the cable.

6.3.3.2 Run, test, verify and correct

After calibration, the system has to be stressed to verify that the given values are consistent with the current conditions. Several methods (detailed hereafter) are available for such purposes.

6.3.3.2.1 Getting the data as expected?

The corrected values (i.e. after the calibration correction, see Chapter 7) should be consistent with the real conditions, while considering the uncertainty. A few questions can help the reader to realize those tests. Does the system provide all the expected data? Are the zeros (water level at zero, velocity at zero, no rain) properly measured? Is it possible to create artificial events (rain, discharge in a pipe) to verify the values given by the system?

6.3.3.2.2 Redundancy, tracing experiment

If it is rather easy to verify the zeros for each sensor, the creation of artificial values (rain, discharge, water level) is slightly more complicated but achievable: injection of known discharge, simulation of different

Table 6.7 Recommended methods to check the data recorded by the measuring system (except for zeros).

Measurements	Technology	Recommended methods
Rain	• Rain gauges • Radar • Others	• Mariotte bottle or calibrated pump • Comparison with rain gauges (if available) • Comparison with rain gauges (if available)
Water level	• US or optical measurements • Pressure	• Horizontal plate under the sensors • Automatic restart procedures
Velocity	• All technologies	• Velocity at the free surface
Discharges	• All technologies	• Cumulated volumes • Mass balance • Tracing experiment

water levels with pumps or fire hydrants, known rain event with a Mariotte bottle, etc. The value recorded in the final file by each sensor should be consistent with the generated artificial value. Table 6.7 suggests some methods to verify recorded values.

Once each sensor has been independently verified, the consistency between each other (if available) must be checked, i.e. the test of the comparison between pairs or sets of values must be applied to all redundant (e.g. two water levels sensors) or correlated (e.g. water level and flow velocity for flow without backwater effects) sensors. The potential difference(s) between the values given by different sensors or calculated by a combination of them must be investigated: installation errors, defaults during the calibration procedures, site or hydraulic condition effects, errors in settings of data loggers, unit conversion, etc. Hypotheses should be tested and verified while avoiding tuning some magic parameters to get consistent values. It is important to provide evidence of proper functioning based on a series of tests rather than simply hoping that the system will work as expected or designed. Proper qualification of the monitoring system is thus fundamental.

6.3.3.2.3 Verify with data

If the system to be verified has several monitoring stations or locations (like inlet, outlet and volume of storage tank or flowmeter in pipe downstream of a pumping station), or several sensors to measure the same value at the same place, other tests can be performed: mass balance at the inlet/outlet of a tank, comparison of cumulated volume (to identify potential drifts), discharges between two points without connection, etc. The design of such tests is necessarily site dependent. There are numerous tests to achieve such comparisons (see Chapter 9 on data validation). A single piece of advice: question, be critical and challenging, and be impartial regarding the design that has just been built.

6.3.3.3 Conclusion after the first tests

After all qualification tests are run (calibration and verification) and assuming the station successfully passed them, the station is ready to be used as designed. Otherwise, understanding the reasons for errors, failures and disagreements with initial design is an important and necessary phase until all corrections are made.

6.3.4 Once the monitoring station is operational

The previously proposed tests can be applied periodically to ensure the system is working fine, even independently from the highly recommended data validation procedures. Two other recommendations should be followed to ensure the right use of the facilities: a site book, to keep track of all the events that occur (expected or unexpected) and continuous training of the employees operating the station.

6.3.4.1 Site book

Derived from a rather old-fashion laboratory book, a physical or digital site book is useful to keep track of observations made by the user. Both versions have pros and cons: a physical book is easier to set up, and efficient for writing down the observations. However, this type of book can be lost, destroyed (e.g. by flooding), and does not allow recording of pictures/photos. The two versions used at the same site can combine the advantages of both, while leading to the need of having to look at both versions. All kind of events should be noted in such a tool: defect of components, misfunctioning, abnormal discharges/events, strange smells, etc.: the list is unlimited. Those data may seem useless at first glance, but they can be helpful for later understanding of misfunctioning, improvements of future designs, and data validation. Some key points need to be written down: names of the people who recorded the observation, date, hour, duration, observations (direct or indirect), reasons or hypotheses to explain the observations (see Table 6.8). The notion of direct and indirect observations needs to be explicit. As an example, an abnormally high discharge can be directly (e.g. the observers can see the flow) or indirectly (e.g. noise in the pipe, new floating materials dumped at high elevation) observed. Indirect observations require more details in the section 'reasons or hypotheses'.

There are numerous solutions available for a digital site book: e.g. SharePoint, file exchange systems, an e-mail address for each station or an electronic calendar. Each solution presents pros and cons, and the selected solution should fit everyone: from the technician in charge of the daily maintenance to the manager.

6.3.4.2 Continuous training of involved people

The importance of continuous training is often underestimated by managers. This is the only way to keep every employee involved in measurements updated: technology progresses fast, new measuring methods or data transmission protocols are available, and standards change. This technological watchfulness requires few actions: exchange between practitioners, visits of/by manufacturers, workshops, readings, internal and external trainings, etc. and permanent questioning.

Table 6.8 Example of content in a physical site book.

Date	29/02/2016
Hour	19:07 UT
Observer	Iris Pear & JP Manoeuvre
Observations	Inconsistencies between water level on the transmitter screens
Action taken	Verification of each sensor with horizontal plate
Identified reason	Obstacle under the US sensor
Problem solved?	Yes
When did the problem start?	I don't know (to be checked with the data)
Additional info	Picture sent by e-mail to the manager (29/02/2016, 19h33 UT)
Suggestion to avoid this problem occurring again	Installation of a grid with large mesh upstream of the measuring location

Figure 6.14 Photo of the leaping weir in Chassieu, France. *Source*: Nicolas Walcker (INSA Lyon).

6.3.5 Example of micro design

Measuring discharges in large pipes is rather challenging: the wide range of hydraulic conditions often requires different technologies to ensure good measurements. The Chassieu OTHU site in Lyon, France has been monitored since the early 2000s. The inlet of the stormwater settling tank in Chassieu is a concrete pipe (circular, diameter of 1600 mm, slope of 1%). Some velocity and water level sensors have been placed in the pipe to measure discharges.

After 10 years of use, feedbacks highlight rather doubtful values when the water level is below 7 cm, due to the disturbances created by the Doppler probes on the measured velocities in their close surroundings. In order to improve the knowledge on the dynamics of small rain events, a leaping-weir has been installed at the pipe outlet, downstream of the previously installed sensors: when the discharge is lower than 4 L/s in the pipe, all the flow passes through the leaping weir (Figure 6.14) and is measured by an accurate electromagnetic flowmeter in a pressurized pipe. For higher discharges, the existing set-up (ultra sound and pressure sensors coupled to flow velocity Doppler sensors) is used, since a part of the flow goes straight, without passing through the electromagnetic flowmeter.

6.4 ADVANCED AND EMERGING MONITORING TECHNOLOGIES

This section briefly introduces some emerging techniques, mainly used in research and sometimes only in the laboratory: the authors lack comprehensive feedback on these techniques. However, they seem promising and once more and robust experience is obtained they will be available in future literature reviews.

6.4.1 Event detection

A recent study introduced a rather robust and simple method to detect overflow events and record their duration (Hofer *et al.*, 2018). This method requires two thermometers: one in the main pipe and one at the invert level of the overflow pipe: once the temperatures are equal at both sites, there is an overflow. As soon as temperatures diverge again, while taking into account the thermal inertia of both sensors, the overflow is finished.

6.4.2 DTS for infiltration

Over the past decade the application of distributed temperature sensing (DTS) using glass fibre cables to detect infiltration in sewer systems or the presence of misconnections (stormwater discharges into wastewater systems or vice-versa) has gained popularity (Hoes *et al.*, 2009; Nienhuis *et al.*, 2013). The measuring principle is described in Section 3.5.3.1.8. In practice there are several issues to reckon with when planning to perform a DTS measurement. A first and very important issue is the attenuation of the optical signal when using long cables, especially the signal loss in connectors between cables is a point for attention (Tyler *et al.*, 2009). In most practical cases one needs to use a number of cables that have to be connected (Figure 6.15) to cover the whole length of the sewer reach one wishes to monitor, typically some kilometres. The reason for this is that inserting huge lengths (in the order of 1–2 kilometres) of cables asks for the deployment of a relatively large team of workers and the availability of heavy equipment. In addition, the risk of a cable getting stuck or broken increases with the length of the cable (Schilperoort *et al.*, 2016). A trade-off has to be made between reducing the number of connectors and the length of the cables used. In this sense the quality of the connectors improved over time allowing for use of shorter cables. At present the application of 200 m long cables is possible given the strongly increased quality of optical connectors applicable under field conditions. These cables are relatively easy to handle (reduced weight) and allow for relatively simple adaptations in the sewer reach covered during the measuring campaign, as only one or a few cable parts have to be replaced.

Further, one is advised to plan the sewer reach well in advance; correcting the route of an optical cable once installed is a tedious and costly task. In most cases the duration of the monitoring campaign is in the order of weeks to months. Therefore, in planning the route, make sure there is a location where a container holding the measuring computer can be stationed, while accounting for traffic and operator safety, and the availability of power supply and preferably data-communication facilities. When choosing the route, make sure there are enough locations at which the cable can be accessed for temperature calibration.

Figure 6.15 Connecting cable segments using an optical connector for DTS measurements. *Source*: courtesy Remy Schilperoort (Partners4UrbanWater).

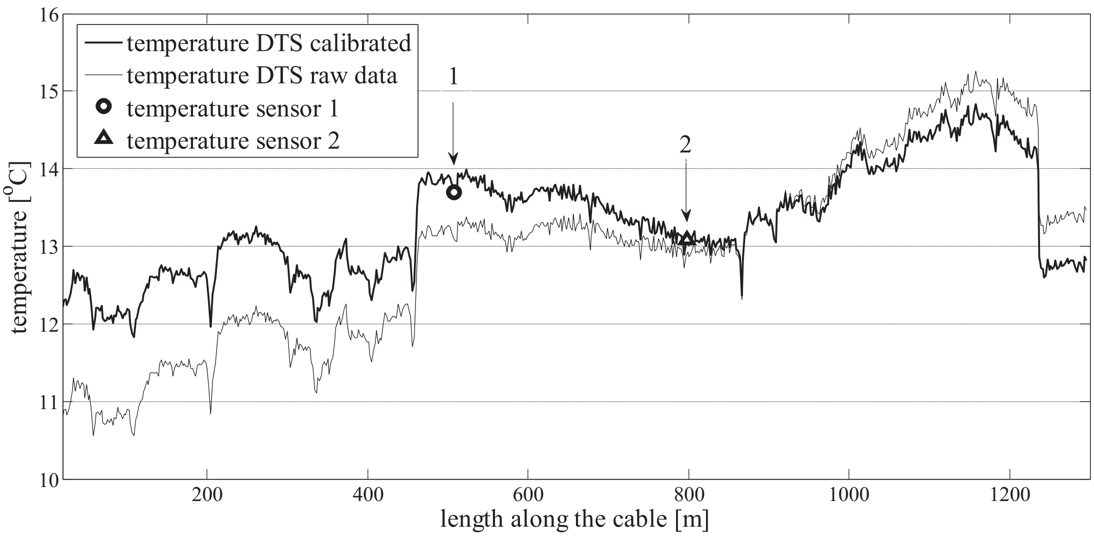

Figure 6.16 Example of a calibration read-out of a DTS cable, along with the corrected result. *Source*: courtesy Remy Schilperoort (Partners4UrbanWater).

For calibration, the temperature is measured at some points along the cable using an accurate thermometer (Pt 100) or, alternatively one can insert parts of the cable (at known distances from the measuring computer) in a bucket of melting ice. Figure 6.16 shows an example of a calibration result.

In most cases one is not really interested in the absolute temperature but in the change over time and its location, making an accurate calibration for temperature values unnecessary. However, one still needs to verify the indicated distances. This can be done by exposing the cable at known locations to a low (or) high temperature, e.g. by simply pouring water with a temperature that is likely to differ from the water present in the system onto the cable (Figure 6.17).

6.4.3 Optical methods for determining flow velocity fields

The application of optical based methods for determination of flow velocity fields in urban drainage has been applied for about the last decade (e.g. Jeanbourquin *et al.*, 2011). LS-PIV (large scale particle image velocimetry) and LS-PTV (large scale particle tracking velocimetry) are non-invasive technologies for determining flow velocity fields. This type of method is becoming more and more popular and emerged from a range of optical measuring methods developed and applied in the experimental study of fluid flow fields since the 1980–1990s. Without going into details of the technologies, a brief description is given hereafter.

Using video footage of floating objects on the surface of a river, channel or conduit in which small light-reflecting or light-emitting objects are present (either introduced as a seeding with known, tailored properties or naturally present), the displacement of these objects between two successive frames is determined and forms the basis for determining a local velocity vector. The difference between PIV and PTV is that the former is based on correlation between successive images, while the latter is based on the movement of individual particles. For a measuring set-up in the field, these differences are unimportant as the differences are mainly found in the post-processing of the raw data (images/video-footage). In

Figure 6.17 Pouring cold water onto the cable, in the underground sewer, using a hose and a funnel for validation of the measured distances by the DTS system. *Source*: courtesy Remy Schilperoort (Partners4UrbanWater).

Figure 6.18 an example of a (field) result from LS-PTV measurement in a wastewater pumping station is shown.

An important issue is ensuring the right light conditions. There is no general rule for this, the settings of light intensity, aperture and exposure time need to be determined by a trial-and-error process. There are however some practical limitations when doing so:

- The exposure time should be relatively short to avoid motion blur as this reduces the useability of the footage for post-processing.
- Depending on the position and orientation of the camera, the aperture has to be chosen in such a manner that the whole region of interest is within focus, this may limit the exposure time.
- The time interval between two frames has to be known and must be significantly larger than the exposure time and in such a magnitude that between frames a noticeable change in position of the markers/seedings has occurred. When observing a pumping station under dry weather conditions, a time interval of 60 seconds between frames is sufficient, while when observing flow in a channel during a storm event, a framerate of 30 frames per second or more may be needed. An optimal, or at least a robust, combination of exposure time, framerate, light conditions and image post-processing settings has to be determined by trial and error.

For an in-depth treatise on PIV, LS-PIV and (LS-)PTV the reader is referred to literature (see e.g. Adrian & Westerweel, 2010; Le Coz *et al.*, 2010).

For practical applications, the following issues have to be considered:

- Camera lenses suffer from optical distortion. To avoid systematic deviations in the result this should be compensated for. Heikkila & Silvén (1997) describe a method to do so in some detail, which is implemented in e.g. a Matlab® function.

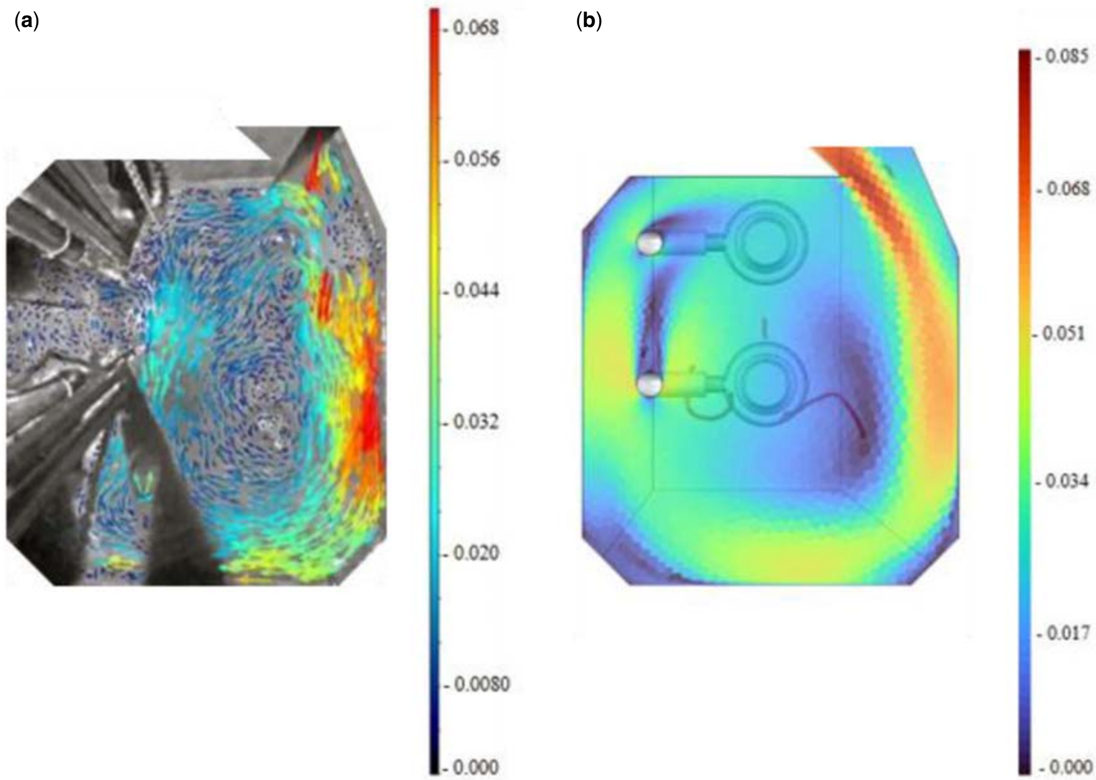

Figure 6.18 Measured and CFD (computational fluid dynamics) modelled surface flow velocities in a pump sump: (a) measured velocities; (b) CFD computed velocities. *Source*: courtesy Alex Duinmeijer (Municipality Rotterdam/TU Delft), based on Duinmeijer (2020).

- There must be a quantified link between pixel positions in the camera's sensor and the real world in 3D. This implies that the exact locations of a number (at least four, but preferably more) of reference points in the field of view (FOV) of the camera have to be known. These reference points should preferably be in, or in the direct vicinity of, the plane on which the camera focusses.
- Make sure the field of interest is covered by the camera's FOV after mounting it.
- In underground spaces, an external light source is needed. Safety issues, especially ATEX requirements (see Chapter 7) are a matter of concern.
- The camera should be mounted in such a manner that its position is and remains fixed throughout the whole length of the monitoring campaign. One is well advised to recheck the reproduction of the position of known elements (reference points) in the image directly after a site visit for maintenance purposes, or to do so on a regular basis e.g. at least prior to each post-processing session.
- Especially in wastewater systems, corrosion is often an issue. Therefore, make sure the equipment is mounted in a protective casing (Figure 6.19).
- When applying LS-PIV or LS-PTV in the field, make sure the effect of wind on the floating markers is minimal, as this can result in significant systematic errors in the end-result.
- Open-source software for post-processing and additional information on LS-PIV can be found in INRAE (2020).

(a) (b) (c)

Figure 6.19 Example of cameras installed in a wastewater pumping station: (a) Camera including electronics; (b) Camera housing, (c) Two cameras installed at a grate in a wastewater pumping station. *Source*: courtesy Antonio Moreno-Rodenas (Deltares).

6.5 SUMMARY AND TRANSITION

In this chapter, macro and micro design of measuring networks and measuring set-ups are discussed, as they are a mixture of theoretical considerations and a compilation of practical experience. This topic is prone to rapid development. It is and remains a challenge to communicate developments and share practical experience. In this respect, readers, scientists and practitioners alike, are invited to share their finding through (inter)national platforms, e.g. through working groups (e.g. the IWA/IAHR WG on metrology in Urban Drainage). Key to making a sound macro design of a monitoring network is that all knowledge available about the network should be used, be it model results, historic observations on e.g. flooding events and/or citizens' complaints. The latter sources provide important clues on 'weak spots' in the system that are locations where additional information on the processes taking place are of importance for the systems manager. In developing the macro design some aspects of the micro design are being decided upon, e.g. when the availability of infrastructure like power supply is not taken in account, this automatically implies that equipment needs to have its own power supply unit. These choices in turn have consequences on the operational phase which is discussed in Chapter 7 on operation and maintenance.

REFERENCES

Adrian R. J. & Westerweel J. (2010). *Particle Image Velocimetry*. Cambridge University Press, New York (USA), 586 p. ISBN 978-0521440080.

Blumensaat F., Dicht S. & Ebi C. (2019). Niedrigenergiefunk im Untergrund - Möglichkeiten und Grenzen einer neuen Datenfernübertragung in der Siedlungsentwässerung [Underground low energy radio communication - Possibilities and limitations for data acquisition in wastewater systems]. *Aqua & Gas*, **99**(3), 52–60. (in German).

Blumensaat F., Scheidegger A., Ebi C., Dicht S., Schaltegger F., Rüst A., Schmitt U. & Maurer M. (2018). Digitalization meets reality – Concept and experiences from long-term wireless data collection with 50+ sewer monitors. *Proceedings of the 11th International Conference on Urban Drainage Modelling*, 23–26 Sept., Palermo, Italy, 360–363.

Boomgaard M. E., Langeveld J. G. & Clemens F. H. L. R. (2001). Wastewater system optimisation using genetic algorithms. *Proceedings of the Specialty Symposium on Urban Drainage Modeling, at the World Water and Environmental Resources Congress*, May 20–24, Orlando, Florida, United States, 788–796. doi: 10.1061/40583(275)74.

Bruni G., Reinoso R., van de Giessen N. C., Clemens F. H. L. R. & ten Veldhuis J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth Systems Sciences*, **19**(2), 691–709. doi: 10.5194/hess-19-691-2015.

Chacon-Hurtado J., Alfonso L. & Solomatine D. (2017). Rainfall and streamflow sensor network design: a review of applications, classification, and a proposed framework. *Hydrology and Earth System Sciences*, **21**(6), 3071–3091. doi: 10.5194/hess-21-3071-2017.

Cherqui F., James R., Poelsma P., Burns M. J., Szota C., Fletcher T. & Bertrand-Krajewski J.-L. (2020). A platform and protocol to standardise the test and selection of low-cos sensors for water level monitoring. *H2Open Journal*, **3**(1), 437–456. doi: 10.2166/h2oj.2020.050.

CIWEM (2017). *Urban Drainage Group - Code of Practice for the Hydraulic Modelling of Urban Drainage*. CIWEM - Chartered Institution of Water and Environmental Management, London (UK), 169 p. Available at https://www.ciwem.org/assets/pdf/Special%20Interest%20Groups/Urban%20Drainage%20Group/Code%20of%20Practice%20for%20the%20Hydraulic%20Modelling%20of%20Ur.pdf (accessed 05 Dec. 2020).

Clemens F. H. L. R. (2001). *Hydrodynamic Model in Urban Drainage: Calibration and Application*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 388 p. ISBN 90-407-2163-7.

Desimoni E. & Brunetti B. (2015). About estimating the limit of detection by the signal to noise approach. *Pharmaceutica Analytica Acta*, **6**(3), 4 p. doi: 10.4172/2153-2435.1000355.

Diaz-Delgadillo F. J., Montiel O. & Sepulveda R. (2016). Reducing the size of traveling salesman problems using vaccination by fuzzy selector. *Expert Systems with Applications*, **49**, 20–30. doi: 10.1016/j.eswa.2015.11.026.

Duinmeijer S. P. A. (2020). *On the Free-Surface Vortex Driven Motion of Buoyant Particles*. PhD thesis, Delft University of Technology, Delft, The Netherlands, 176 p. ISBN 978-94-6366-271-0.

Ebi C., Schaltegger F., Rüst A. & Blumensaat F. (2019). Synchronous LoRa mesh network to monitor processes in underground infrastructure. *IEEE Access*, **7**, 57663–57677. doi: 10.1109/ACCESS.2019.2913985.

Fencl M., Rieckermann J., Sykora P., Stransky D. & Bares V. (2015). Commercial microwave links instead of rain gauges: fiction or reality? *Water Science and Technology*, **71**(1), 31–37. doi: 10.2166/wst.2014.466.

Fu Y., Jing C. & Du M. (2016). Urban rain gauge siting selection based on GIS-multicriteria analysis. *The International archives of the Photogrammetry, Remote Sensing and Spatial Sciences, Volume XLI-B2, XXIII ISPRS Congress*, 12–19 July, Prague, Czech Republic, 185–190. doi: 10.5194/isprsarchives-XLI-B2-185-2016.

Grafakos L. (2014). *Classical Fourier Analysis, 3rd edn*. Springer, New York (USA), 655 p. ISBN 978-1-4939-1193-6. doi: 10.1007/978-1-4939-1194-3.

Heikkila J. & Silvén O. (1997). A four-step camera calibration procedure with implicit image correction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 17–19 June, San Juan, Porto Rico, USA, 1106–1112. doi: 10.1109/CVPR.1997.609468.

Henckens G. J. R. & Clemens F. H. L. R. (2004). Design and optimisation of monitoring networks in urban drainage. *Proceedings of the 19th European Junior Scientists Workshop on "Process Data and Integrated Urban Water Modelling"*, 11–14 March, Meaux-la-Montagne, France, 8 p.

Hoes O. A. C., Schilperoort R. P. S., Luxemburg W. M. J., Clemens F. H. L. R. & van de Giessen N. C. (2009). Locating illicit connections in storm water sewers using fiber optic distributed temperature sensing. *Water Research*, **43**(20), 5187–5197. doi: 10.1016/j.watres.2009.08.020.

Hofer T., Montserrat A., Gruber G., Gamerith V., Corominas L. & Muschalla D. (2018). A robust and accurate surrogate method for monitoring the frequency and duration of combined sewer overflows. *Environmental Monitoring and Assessment*, **190**, 209, 18 p. doi: 10.1007/s10661-018-6589-3.

INRAE (2020). Fudaa-LSPIV Information Webpage. Available at https://riverhydraulics.inrae.fr/en/tools/measurement-software/fudaa-lspiv-2/ (accessed 04 Dec. 2020).

Jeanbourquin D., Sage D., Nguyen L., Schaeli B., Kayal S., Barry D. A. & Rossi L. (2011). Flow measurements in sewer systems based on image analysis: automatic flow velocity algorithm. *Water Science and Technology*, **64**(5), 1108–1114. doi: 10.2166/wst.2011.176.

Kapelan Z., Savic D. A. & Walters G. A. (2003). Multi-objective sampling design for water distribution model calibration. *Journal of Water Resources Planning and Management*, **129**(6), 466–479. doi: 10.1061/(ASCE)0733-9496(2003)129:6(466).

Kapelan Z., Savic D. A. & Walters G. A. (2005). Optimal sampling design methodologies for water distribution model calibration. *Journal of Hydraulic Engineering*, **131**(3), 190–200. doi: 10.1061/(ASCE)0733-9429(2005)131:3(190).

Le Coz J., Hauet A., Pierrefeu G., Dramais G. & Camenen B. (2010). Performance of image-based velocimetry (LSPIV) applied to flash-flood discharge measurements in Mediterranean rivers. *Journal of Hydrology*, **394**(1–2), 42–52. doi: 10.1016/j.jhydrol.2010.05.049.

Lei J. & Schilling W. (1993). Requirements of spatial rain data resolution in urban runoff simulation. *Proceedings of the 6th ICUD – International Conference on Urban Drainage*, 13–17 Sept., Niagara Falls, Canada, Vol. 1, 447–452.

Lepot M. (2012). *Mesurage en continu des flux polluants en MES et DCO en réseau d'assainissement [Continuous monitoring of TSS and COD fluxes in sewers]*. PhD thesis, INSA Lyon, Villeurbanne, France, 257 p. (in French).

Lepot M., Aubin J.-B. & Clemens F. (2017). Interpolation in time series: an introductive overview of existing methods, their performance criteria and uncertainty assessment. *Water*, **9**(10), 796, 20 p. doi: 10.3390/w9100796.

Mikulak R. J., McDermott R. E. & Beauregard M. R. (2008). *The Basics of FMEA, 2nd edn*. Productivity Press, New York (USA), 90 p. ISBN 978-1563273773.

Nienhuis J., de Haan C., Langeveld J., Klootwijk M. & Clemens F. (2013). Assessment of detection limits of fiber-optic distributed temperature sensing for detection of illicit connections. *Water Science and Technology*, **67**(12), 2712–2718. doi: 10.2166/wst.2013.176.

Ochoa-Rodriguez S., Wang L.-P., Gires A., Pina R. D., Reinoso-Rondinel R., Bruni G., Ichiba A., Gaitan S., Cristiano E., van Assel J., Kroll S., Murlà-Tuyls D., Tisserand B., Schertzer D., Tchiguirinskaia I., Ono C., Willems P. & ten Veldhuis M.-C. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*, **531**, 389–407. doi: 10.1016/j.jhydrol.2015.05.035.

Ochoa-Rodriguez S., Wang L.-P., Willems P. & Onof C. (2019). A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resources Research*, **55**(8), 6356–6391. doi: 10.1029/2018WR023332.

Orfeo D., Burns D., Farrell R., Qin M., Mitchell H., Ou C., Xia T. & Huston D. R. (2018). Mechano-magnetic telemetry for underground water infrastructure monitoring. *Frontiers in Built Environment*, **4**, 29, 14 p. doi: 10.3389/fbuil.2018.00029.

Raspe R. E. (1895). *The Surprising Adventures of Baron Munchausen*. Available at https://www.gutenberg.org/files/3154/3154-h/3154-h.htm (accessed 05 Dec. 2020).

Rodríguez-Iturbe I. & Mejía J. M. (1974). The design of rainfall networks in time and space. *Water Resources Research*, **10**(4), 713–728. doi: 10.1029/WR010i004p00713.

Ruiz-Cardenas R., Ferreira M. A. R. & Schmidt A. M. (2010). Stochastic search algorithms for optimal design of monitoring networks. *Environmetrics*, **21**(1), 102–112. doi: 10.1002/env.989.

Schellart A. N. A., Wang L. & Onof C. (2012). High resolution measurement and analysis in a small urban catchment. *Proceedings of the 9th International Workshop on Precipitation in Urban Areas: Urban Challenges in Rainfall Analysis*, 6–9 Dec., St. Moritz, Switzerland, 115–120. ISBN 9783906031217.

Schilperoort R. P. S., de Haan C., Liefting E., Nijhof H., Hof A., Meijer C. & Langeveld J. (2016). Advances in practical applicability of DTS in sewer systems. *Proceedings of the 8th International Conference on Sewer Processes and Networks*, 31 Aug.–2 Sept., Rotterdam, The Netherlands, 33–37.

Schilperoort R. P. S., Dirksen J. & Clemens F. H. L.R. (2008). Practical aspects for long-term monitoring campaigns: pitfalls to avoid to maximize data yield. *Proceedings of the 11th International Conference on Urban Drainage*, 31 Aug.–5 Sept., Edinburgh, Scotland, UK, 10 p.

Schilperoort R. P. S., Dirksen J., Langeveld J. G. & Clemens F. H. L. R. (2012). Assessing characteristic time and space scales of in-sewer processes by analysis of one year of continuous in-sewer monitoring data. *Water Science and Technology*, **66**(8), 1640–1620. doi: 10.2166/wst.2012.115.

Shaghaghian M. R. & Abedini M. J. (2013). Rain gauge network design using coupled geostatistical and multivariate techniques. *Scientia Iranica*, **20**(2), 259–269. doi: 10.1016/j.scient.2012.11.014.

Sivakumar B. (2010). Fractal analysis of rainfall observed in two different climatic regions. *Hydrological Sciences Journal*, **45**(5), 727–738. doi: 10.1080/02626660009492373.

Thompson K. E., Vamvakeridou-Lyroudia L., Kapelan Z. & Savic D. (2011). *Optimal macro-location methods for sensor placement in urban water systems – Literature review*. EU FP7 project PREPARED, report 2011.004, 75 p. Available at https://emps.exeter.ac.uk/media/universityofexeter/emps/research/cws/downloads/PREPARED_Deliverable_3_5_1_final.pdf (accessed 05 Dec. 2020).

Tiwari S., Jha S. K. & Singh A. (2020). Quantification of node importance in rain gauge network: influence of temporal resolution and rain gauge density. *Scientific Reports*, **10**(1), 9761, 17 p. doi: 10.1038/s41598-020-66363-5.

Tscheikner-Gratl F., Bellos V., Schellart A., Moreno-Rodenas A., Muthusamy M., Langeveld J., Clemens F., Benedetti L., Rico-Ranires M. A., Fernandez de Carvalho R., Breuer L., Shucksmith J., Heuvelink G. B. M. & Tait S. (2019). Recent insights on uncertainties present in integrated catchment water quality modelling. *Water Research*, **150**, 368–379. doi: 10.1016/j.watres.2018.11.079.

Tyler S. W., Selker J. S., Hausner M. B., Hatch C. E., Torgersen T., Thodal C. E. & Schladow S. G. (2009). Environmental temperature sensing using Raman spectra DTS fiber-optic methods. *Water Resources Research*, **45**, W00D23, 11 p. doi: 10.1029/2008WR007052.

van Luijtelaar H. & van Rebergen E. W. (1997). Guidelines for hydrodynamic calculations on urban drainage in The Netherlands: backgrounds and examples. *Water Science and Technology*, **36**(8–9), 253–258. doi: 10.1016/S0273(97)00590-8.

Walcker N., Bertrand-Krajewski J.-L., Vacherie S., Lepot M., Castebrunet H., Barraud S. & Lipeme Kouyi G. (2018). Une nouvelle station de mesure pour l'acquisition de séries chronologiques en hydrologie urbaine [A new monitoring station for the acquisition of time series in urban hydrology]. *Techniques Sciences Méthodes*, **3**, 55–64. doi: 10.1051/tsm/201803055. (in French).

WMO (2018). *Guide to instruments and methods of observation - Volume I: Measurement of meteorological variables*. World Meteorological Organization, Geneva (Switzerland), 573 p. ISBN 978-92-63-10008-5. Available at https://library.wmo.int/doc_num.php?explnum_id=10179 (accessed 05 Dec. 2020).