

## Discussion note about system identification

The output  $\bar{p}_K$  is given by *Equation: (1)*, which is a static equation

$$\bar{p}_K = K^T \bar{H}_T^{-T} f_T(A_2^T q_C + A_3 K \bar{d}_K - A_3 D v_D \sigma) - \underbrace{K^T \bar{H}_T^{-T} \hat{H}_T^T (\hat{p} + \hat{h})}_{\text{Linear term}} - K^T \bar{h}. \quad (1)$$

We consider it as an input-output model where the inputs:

$K^T \bar{h} = \bar{h}_K$	is the elevation of inlets ,
$\hat{p}$	is the pressure in the WTs,
$\hat{h}$	is the elevation of the WTs,
$\sigma$	is the total demand,
$\bar{d}_K$	is the inlet flows,
$q_C$	is the flows in $\mathcal{C}$ .

and the output is  $\bar{p}_K$ , which is the vector of inlet pressures. If we consider this input-output relation as a black-box model, all inputs and outputs are known and can be extracted (almost) directly from EPANET, except  $q_C$  for which we have the constraint. For the identification, Radial Basis Function(RBF) Networks can be used, assuming that we do not have any priori information about the system structure.

For the dynamics, some non-recursive Neural Network(NN) based approach can be used, however the dynamics are also in terms of  $q_C$ .

$$\Lambda \dot{\hat{p}} = -(\hat{H}_C - \hat{H}_T \bar{H}_T^{-1} \bar{H}_C) q_C - \hat{H}_T \bar{H}_T^{-1} K \bar{d}_K + \hat{H}_T \bar{H}_T^{-1} D v_D \sigma. \quad (2)$$

The dynamics are linear, therefore some other linear black-box model methods are also suitable.

The constraint on  $q_C$  is given by *Equation: (3)*

$$f_C(q_C) - A_1(\hat{p} + \hat{h}) + A_2 f_T(A_2^T q_C + A_3 K \bar{d}_K - A_3 D v_D \sigma) = 0. \quad (3)$$

where

$$\begin{aligned} A_1 &= \hat{H}_C^T - \bar{H}_C^T \bar{H}_T^{-T} \hat{H}_T^T, \\ A_2 &= -\bar{H}_T^{-1} \bar{H}_C, \\ A_3 &= \bar{H}_T^{-1}. \end{aligned}$$

This is an implicit function and  $q_C$  is unknown. Furthermore,  $q_C$  is defined by the partitioning of the network, therefore the only way to extract this data from EPANET is by defining the set  $\mathcal{C}$  first. In order to solve for  $q_C$ , the structure of the system is considered and partitioned. The data used( $\bar{d}_K, \hat{p}, \hat{h}, \sigma$ ) are the same as for the models regarding the dynamics and the outputs.

$q_C$  is calculated using the system matrices for a 48 hours period, and then it can be used as the input signal for the dynamics and output model.

*Q1:* As there is no measurement data on  $q_C$ , it is calculated by using the measured inputs and using the system model. (The model is used, therefore the constraint is not treated as black-box.) Is the idea here a possible way to continue or should the constraints on  $q_C$  handled differently?

*Q2:* In *Equation: (1)*, the linear terms can be separated from the non-linear part if we use the system matrices.  $\hat{h}$  and  $\bar{h}_K$  are constants, therefore they are treated as an offset. They are subtracted from the measured  $\bar{p}_K$  inlet pressures.  $\bar{p}$  is measured, therefore the linear term regarding  $\bar{p}$  is also subtracted. When it is done, the identification is done on the non-linear terms, including only the remaining inputs ( $\sigma, \bar{d}_K, q_C$ ).