DOI: 10.31577/ahs-2023-0024.02.0030

ACTA HYDROLOGICA SLOVACA

Volume 24, No. 2, 2023, 275 – 284

Numerical tests and sensitivity analysis of pollution source localisation tool applied on open channel system

Yvetta VELÍSKOVÁ*, Marek SOKÁČ, Maryam BARATI MOGHADDAM

This paper is focused on solving the inverse problem of pollution spreading in open channel system in order to find the location of pollutant sources. The goal was to propose localisation procedure, which should be operative and fast enough to enable quick interventions to prevent the spread of pollution. The proposed method, as well as the overall localisation procedure was numerically tested on a real sewer system data, which represents in this case an extensive open channel network structure with free surface flow. As part of the numerical tests of the model, a sensitivity analysis was also carried out. The sensitivity analysis was focused on the resulting localization error depending on concentration measurement errors. The numerical test results were successful and confirmed applicability of proposed localization tool in real conditions, the sensitivity analysis points to the influence of individual parameters on localization accuracy.

KEY WORDS: inverse problem, source localisation, pollution source, sensitivity analysis, numerical test

Introduction

Population growth and anthropogenic activities cause many management challenges (Novák, 2022), especially in secure disposal of water for various uses and protect water resources and environment from contamination, especially harmful substances such as emerging organic contaminants, heavy metals, biodegradable pollutants, micropollutants, pesticides, industrial chemicals, petroleum hydrocarbons and so on. Exposure to these contaminants cause serious health risks and environmental problems. Moreover, illegal discharges of high volumes of hazardous materials from industrial complexes may impair the function of the wastewater treatment plants, in case of pollution outflow in agricultural area or field can endanger the quality of the water body (Kováčová, 2023). In most cases the source of pollutant discharge is unknown and due to the complex factors and conditions in environment, it is often very difficult to identify the exact spill location. Hence, development of methods that are able to track down contaminant sources location can help in minimizing the adverse impacts as well as preventing future illegal spills.

Source identification is an inverse problem of the transport equation. From a mathematical point of view, contaminant source localization is an ill posed problem does not fulfil the well-posedness criteria of Hadamard (1923). Based on Hadamard's definition, a problem is well-posed if its solution is existent, unique and stable (that is the continuous dependence of the solution to input data). A problem which lacks any of these features is called an ill-posed problem. The ill-posedness feature of inverse source problem is generally originated from the fact that this problem is under-determined, because only limited observational data are available and the unknown source characteristics should be estimated in a number of instants which is much greater than available observed data. The second reason for the ill-posedness is a direct consequence of the irreversibility of dispersion phenomena, which gradually smooth the pollutant plume and decrease the amount of information that can be obtained from observations (Skaggs and Kabala 1998; Boano et al. 2005).

In the past decades, the problem of identifying unknown pollutant sources has attracted the attention of researchers around the world. Pollution source identification problem has been studied in different environments of watercourses, but this topic has not been addressed in watercourses networks or sewer networks (as limited case of watercourses network) properly. While identifying unknown pollutant sources is one of the essential measures in planning effective management strategies in urban areas. There are only few studies in the literature that dedicated to this topic. Banik et al. (2014) developed a toolkit library for SWMM5 to solve a pollution source identification problem in sewer networks. Their proposed methodology was able to identify the source location as well as its released

concentration, starting time and duration. However, their proposed method needs too much computational effort, especially in the case of large sewer networks, which makes it impractical to apply the method for real life situations. Later, in order to address the issue of excessive computational effort, Banik et al. (2015) presented a pre-screening procedure to apply before the source identification step. The proposed pre-screening procedure identifies a set of candidate nodes using the pollution matrix concept (Kessler, et al. 1998). Results showed that introducing the pre-screening procedure reduces in the required computational time significantly in large networks. It also improves the accuracy of the identification method. However, the effectiveness of the proposed pre-screening procedure is highly dependent to the placement arrangement of sensors. so, it is a factor that should be considered prior to using this method.
Shao et al. (2021) coupled Bayesian inference with SWMM and developed a stochastic source identification model to reconstruct the characteristics of an instantaneous pollutant discharge in sewer networks. Results showed that the performance of proposed method is highly dependent on walking step size, numbers of unknown source parameters and numbers of monitoring sites.
In this paper, we proposed a localization method and tool for identifying pollutant sources in open channel system which can be applied both at natural and urban catchments (watercourses network, sewer network).

Material and methods

Theoretical background

As it was mentioned above, localisation of unknown pollution sources is a challenging task in water resources management. The prevailing majority of existing simulation models is designed to simulate the pollution spreading in the water flow direction only. In current advancing monitoring technologies, a different task can occur: the monitored pollution time course in the cross-flow profile is known, but the pollution source is unknown. The solution of task defined such way is the source determination, which means the source localisation or determination of the pollution time course release parameters (start time, duration, concentration over time) or so-called source intensity function. Solution of such task is often called as solution of the inverse problem.
An inverse problem is the process of calculating the causal factors from a set of observations that produced them. Inverse problems are some of the most important mathematical problems in science and mathematics because they tell us about parameters that we cannot directly observe.
Methods for solving the inverse task related to the area of pollution transport in water environment was presented in the literature several times (Bagtzoglou and Atmadja, 2003; Cheng and Jia, 2010; El Badia and Hamdi, 2007; Ghane et al., 2016; Mazaheri et al., 2015; Telci and Aral, 2011; Wang et al., 2018). However, the research in this field is very often focused mainly on mathematical methods and tests (verification), applications in real conditions are very rare. The problems are mainly the computational demands of the inverse transport models, high uncertainty of the inverse task solution methods as well as the problems connected with the temporal and spatial data collection, sampling, sensitivity, accuracy.
The simplest way how to localize the pollution source is to deploy a number of sensors and to wait until a pollution spill occurs. The occurrence of the investigated pollution presence on particular sensors will determine the area with the pollution source.
Such approach on the other hand requires a large number of sensors and it is extensive time-consuming procedure (the user has to wait until next pollution spills occur, if even occur). Therefore, it is reasonable to develop a methodology for pollution source identification (even though with higher uncertainty) based on a single pollution spill event.
So far, no methods for the inverse task in sewer system were published, only for inverse task solutions in a river system (Ghane et al., 2016), which is hydraulically a similar task, but the main difference is the complexity and size of the area of interest (number of network nodes, branches).
The concept of well- and ill-posedness goes back to (Hadamard, 1923). However, the definition was given by (Courant and Hilbert, 1993). According to the definition, a mathematical problem is called well - posed (in the sense of (Hadamard, 1923) or properly posed or correctly posed if:
1. a solution of the problem exists in a given set of “admissible” solutions (existence condition),
2. the solution is unique (uniqueness condition),
3. the solution depends continuously on input data (stability condition).

If at least one of the three conditions fails, then the problem is called ill-posed (also improperly posed or incorrectly posed problem).
Typically, several approximate solutions can be found in real conditions (locations as well source intensity functions), which fits well with observed data. This is caused by the tree structure of the watercourses system, so the pollution can be transported from the unknown source to the observation point by different ways, whereas the transport time (dispersion process, dilution, time course transformation…) can be very similar, so the resulting concentration time course can be almost identical and by this way the unique solution of the inverse task cannot be ensured.
The source identifiability is closely connected with assumption of some information of the unknown source characteristics. The measured pollution concentration data over time (concentration time course) alone are insufficient source of data for solving the inverse task. Beside the known measured pollution time course, we have unknown the source location as well as the source intensity function. This makes the solution of the inverse task impossible (ill – posed task). To decrease the uncertainty degree of the task, we have to assume

some information about the pollution source. For example, in case of sewer network this could be the instantaneous pollution release, but more realistic would be some specific type of the pollution source intensity function, e.g. the inflow concentration time course. The total mass of the pollution can be estimated based on the measured data (measuring concentration and discharge). For task simplification we assume, that the pollution is conservative; this means, that the pollution is not a subject of chemical or biological degradation process during transport in water. However, for special pollution substances, such a volatile or fast reactive pollution types, this approach should be amended.

Assumptions

The inverse task solutions are usually connected with many unknown characteristics, as well as some methods of their solution require a large number of numerical simulations. Therefore, simplifying assumptions are often used, results of which are simple numerical models (e.g. analytical solutions of partial differential equations) that accelerate the numerical simulations and reduce the computing time required to solve the overall task.

For practical solution of an inverse problem the following assumptions are very often accepted:

1. the monitored substance is conservative,
2. the monitored substance will be not attached to the suspended particles, bed material or sediment, will not react with the other pollutants in water,
3. the monitored substance background concentration will be low (related to the concentration in substance source), or zero, eventually close to zero,
4. confluence points on the watercourses network change the monitored substance concentration, but do not change the concentration time course shape,
5. the monitored substance source is a simple one,
6. knowledge of the source intensity function,
7. the measuring intervals of the detection device will be sufficiently short,
8. the detection limit of the monitoring device - sensor will be sufficiently low,
9. various assumptions adopted in the hydraulic model of the sewer network, which includes:
   a) prismatic channel bed along a reach,
   b) the open channel network will have from the hydraulic point of view a simple tree structure, e.g. there will be no bifurcation (flow split-up) points, no looped circuits in the system structure, the flow path will be unique from each possible pollution point,
   c) there will be no objects (hydraulic obstacles, so called singularities) in the system network, which will completely disturb the pollution time course,
   d) the flow in the open channel system can be regarded as a steady and uniform one, at least during the substance transport from source to the observation point,
   e) it will be possible to determine discharges in each particular sections and related hydraulic and dispersion parameters,
   f) unpredictable and discontinuous pollutant inflows will not occur – e.g. large industrial discharges, varying in time, also flush flood discharges will not occur.

Methods for inverse task solution

There are several approaches to solve the inverse task. According (Barati Moghaddam et al., 2021) the methods can be categorized into three main groups:

1. optimisation methods,
2. stochastic (probabilistic) methods and
3. mathematics based methods and approaches.

The first ones – the optimisation approaches are based on the minimisation of the difference between simulated and observed contaminant concentrations at the observation points by using an optimization algorithm. As the optimisation algorithms can be used "classical" linear programming, least squares regression methods, as well as the "non-classical" methods like the genetic algorithms, artificial neural networks. This approach uses a numeric model for forward simulation and compares the output (modelled) data with the measured data. The numeric model can be an integrated model or an external model, linked through the above-mentioned optimisation procedure. Use of an external model can be very advantageous because of the complexity and reliability of the used model.

The stochastic (probabilistic) methods consider parameters to be random variables rather than fixed deterministic quantities. The major stumbling block to the use of probabilistic models is how to assess the statistical properties of unknown model parameters, considering that there is a large degree of uncertainty in their measured values. (Woodbury et al., 1998)

The last group of the methods solve the inverse problem using direct numerical or analytic methods. These methods are typically very complex and sophisticated, on the other hand less time consuming, more straightforward with more unique results (Barati Moghaddam et al., 2021).

Choosing the approach for the open channel or sewer systems, we focused our effort to the first group of inverse task solution methods, i.e. on optimisation methods.

Optimisation methods are very flexible and relatively easy to create the SW code, eventually the possibility of use of external flow and pollution transport models. Their disadvantage can be the lower level of mathematical apparatus, but regarding to the uncertainties in the whole localisation task it is not reasonable to use high sophisticated mathematic (numeric) methods, which in our opinion will not significantly increase the overall precision of the inverse task solution.

The next disadvantage of proposed method of solving the inverse problem is the necessity to realize a large number of numerical simulations, respectively long calculation time. This disadvantage, however, gradually disappears with the increase in computer technology performance and speed. Our proposed model performs the large number of simulations only once (creating the unit pollutograms), eventually once for specific hydrologic situation. This procedure can be realised in advance, the localization procedure itself is then very quick.

Localisation tool principle and description

The proposed (and developed) localisation tool is based on model creation of so called "unit pollutograms", i.e. on creation of the substance concentration time course for each observation point and for each possible source point (i.e. all network points upstream the observation point). Unit pollutogram means, that the mass of the substance, instantaneous released in the source point is equal to one (arbitrary) unit; the mass of the substance flowing through the observation profile will be the same, i.e. equal to one unit. This modelling principle is very often used in hydraulic (hydrologic) modelling, whereas the unit pollutograms (hydrograms) are often referred as the response function.
The generated unit pollutograms are build up assuming the instantaneous pollution entry (source intensity function, in form of the Dirac impulse input). If a different source intensity function should be applied, the resulting modelled pollutogram in the observation point can be derived simply by the discretisation of the input function to the set of unit instantaneous pollution entries. The resulting (modelled) pollutogram can be simply derived by addition of particular unit pollutograms (response function), eventually multiplicated by the factor according the relative concentration ratio in the discretized source intensity function.
Real pollutogram in this sense means that this pollutogram contains absolute concentration values. This could be easily achieved by multiplying the unit pollutogram by the observed (monitored) mass of the substance, so the amount of the monitored and modelled substance will be equal. The localization procedure in the form of a flowchart is shown on Fig. 1.

Hydraulic model

Hydraulic model is a basic and crucial part of the overall localisation tool. The basic goal of the hydraulic model is to determine the hydraulic conditions in the open channel/sewer network, namely the hydraulic flow rates (discharges) and depending flow velocity for each part (branch) of the system. A basic condition is to have a database, which contains all structural data about the open channel network, e.g. diameters, slope, lengths, coordinates, topology information, etc.
The best option will be wide range monitoring of the necessary hydraulic parameters (velocity, flow rates), but of course, this is not possible, especially in case of wide and large network systems. Typically, nowadays the flow rate measuring devices are stationed only in few points, mainly at the end of a system part. Therefore, it is necessary to use a mathematical model, which will deliver information about the flow rates and other hydraulic parameters across the whole open channel network.

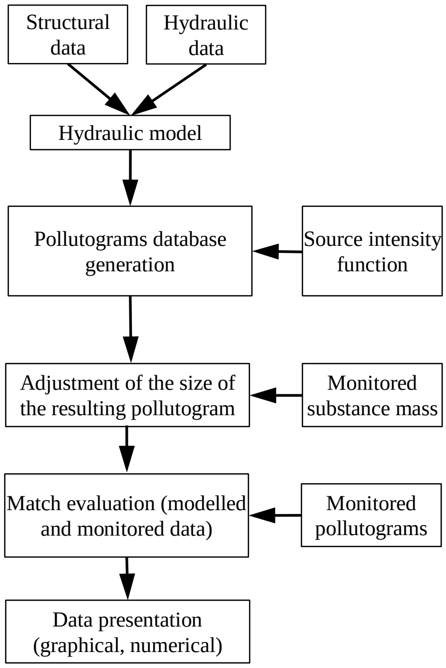

Fig. 1. Flowchart of the localisation tool.

The basic equation for the hydraulic model, used for calculation of the flow rate for specific cross-section of the sewer network is

$$Q_a \frac{N q_s}{86400} + \sum q_p \tag{1}$$

where Q_a is the average flow rate [l s^{-1}], N is the number of inhabitants, connected to the sewer system upstream of the investigated sewer network profile [-], q_s is the specific daily wastewater production per capita [l inh^{-1} day^{-1}], 86400 is the number of seconds per day [sec], the term $\sum q_p$ represents the sum of all point inflows into the sewer system upstream the investigated sewer network profile [l s^{-1}].
Anyway, from this point of view, any monitored information about the flow rates would be very useful. Because of this, implementation of flow in measuring (monitoring) devices (level, velocity sensors) is very welcomed and can significantly improve the localisation procedure precision.

Pollutant transport model

The modelling of the pollutant transport in open channel network consists of two basic and consecutive steps:
1. solution of the pollution transport in a single branch and,
2. confluence and mixing of the substance in junctions (sewer manholes).

For the pollution transport are in the model 5 various methods applied. All of used methods are based on the solution of the advection – dispersion equation (ADE), which describes the transport of solute in flowing water. To simplify an already rather complex task, we used the one-dimensional form ADE. Such simplified ADE has a form (Fischer et al., 1979):

$$\frac{\partial Ac}{\partial t} + \frac{\partial Qc}{\partial x} - \frac{\partial}{\partial x}\left(AD_L \frac{\partial c}{\partial x}\right) = -AKc + C_s\, q \qquad (2)$$

where: c is a mass concentration [g l^{-1}]; D_L is the longitudinal dispersion coefficient [m^2 s^{-1}], A is a discharge area in a cross-section [m^2], Q is a discharge in a sewer [m^3 s^{-1}], K represents a rate of growth or decay of contaminant [s^{-1}], C_s is the concentration of a source, q is a discharge of a source, x is a distance [m].

For a faster response of the proposed localization tool and overall simplification of the procedure, analytical solutions of ADE were used for modelling the forward spread of pollution: Gauss (Fischer et al., 1979; Runkel and Broshears, 1991), Gumbel (Sokáč et al., 2018) and Generalised Extreme Value (GEV) distribution model (Sokáč et al., 2018).

As stated in the theoretical part, it can be assumed that "ideal" (immediate) mixing occurs at the points of confluence (watercourses confluences, manholes on the sewer network). This means that this will change the concentration of the pollutant due to mixing, but the shape of the time course of the concentration of the pollutant (pollutograms) will remain unchanged. Mathematically, mixing at confluence points can be expressed by the following equation (equation of ideal mixing):

$$c(t) = \frac{\sum_{i=1}^{i=N} Q_{i,t}\, c_{i,t}}{\sum_{i=1}^{i=N} Q_{i,t}} \qquad (3)$$

where the $c(t)$ is the concentration in time t [g l^{-1}], $Q_{i,t}$ is the flow rate [m^3 s^{-1}], inflowing in time t from the *i-th* inflow into the confluence point, $c_{i,t}$ is the pollutant concentration [mg l^{-1}] of this inflow in time t, N is the total number of inflow (connected watercourses / sewer branches).

It is obvious, that in case of confluence with one or more lateral inflows with zero pollutant concentration will the resulting (outflow) concentration abruptly decrease; the decrease rate is depending on the mutual ratio of discharges of particular inflows to the confluence node.

Pollutogram database generation

This part of the localisation tool generates the particular pollutograms for each point of the open channel system. The overall process proceeds as follows:

1. generation the substance input in specific point of the network,
2. modelling the pollutogram transformation (i.e. its advection and dispersion) along the adjacent (downstream) branch,
3. record the pollutogram into a database file(s),
4. the process in points 2 and 3 will repeat until the most downstream node of the open channel system is reached,
5. the process in points 1–4 is repeated for each point of the sewer network.

Such procedure will generate the transformed pollutograms between each pair of system nodes where a hydraulic connection between those two nodes.

The pollutograms database generation should be performed for various hydraulic conditions of watercourses system, i.e. for average, as well as for a wider range of flow rates between the minimal and maximal flow rates.

Pollutogram match evaluation – localisation procedure

The last step in the overall localisation procedure is the match evaluation between the modelled pollutograms (stored in the database) and measured pollutograms. This is realized through standard statistical tests, focused on the evaluation of the match of two time courses data series (pollutograms), namely the Root Mean Square Error (RMSE) and the Normalised Root Mean Square Error (NRMSE). The goal of this procedure is to assign to each node on the sewer network a probability, that the source we are looking for will be right at that point. The probability of the source occurrence in particular network points is determined within the value range from 0 to 1 (or 0 – 100%), whereas zero value means there is no possibility, that the source is located in this point; the value one means the highest probability of the source spill occurrence in this point.

It is necessary to known, in which point/location/profile of the open channel system was the substance pollutogram measured, i.e. to know the monitoring profile in the network. As potential pollutant sources are considered all network nodes, located upstream, all downstream nodes will be automatically assigned zero probability. The comparison procedure is performed with all network points located upstream and comparison results between the measured and modelled pollutograms (RMSE, NRMSE) are assigned to particular network points.

Localisation tool includes also option to exclude a part (branch) of the open channel system from the localisation procedure. This option can be used, if there another monitoring device, which does not register the presence of the monitored pollutant ("zero pollutogram") in that network part. This means, that the substance source will be definitely not located upstream of that monitoring device.

In order to compare two series of time (pollutograms), it is necessary first to put both time series on the same basis. In our localisation tool, the mass normalisation was used, i.e., in all cases the total tracer mass corresponds to one dimensionless unit; the time is not normalised. The dimensionless mass normalised concentration is defined:

$$c_d(t) = \frac{Q\, c(t)\, \Delta t}{\sum_{i=0}^{i=\infty} Q\, c_i\, \Delta t} \qquad (4)$$

where $c_d(t)$ is the dimensionless mass normalised concentration [-], c [t] is the substance concentration in time t [kg m^{-3}], Q is the flow rate [m^3 s^{-1}].

The measured and modelled pollutogram, normalised such way can be subsequently subject of the statistical match evaluation, using the above-mentioned statistical methods (RMSE, NRMSE).

The resulting array (network nodes with assigned source location probability) is then sorted to present to the user the nodes with highest source occurrence probability. For better results presentation can be the result array also percentage normalised – the highest probability value can be regarded as 100% probability and the lowest value (typically close to zero or zero) as the 0 % probability. Such normalisation allows to use achieve a full-scale range of the results in graphical interpretation of the results. The normalised values are then assigned to each open channel system node and corresponding data (coordinates, source probability in normalised percentage) for the graphics are exported to a separate file. The graphical results presentation is currently implemented into the AutoCAD graphical software, where the source probability of particular nodes is shown by differentiated by various colour and size of the circles in corresponding open channel system nodes.

Numerical tests of the localisation tool

Since the tests and verification of the localization procedure is not easy in the real conditions of the watercourse system due to the difficulty of obtaining the input data, the localization tool was tested only on numerically generated cases of discharge of a pollutant in a fictitious, simplified sewer network, as a limit case of an open channel system. Input data were chosen on base of a dataset from a real sewer infrastructural database from a town in south-east part of Slovakia. The parameters of the infrastructural database are summarized in Table 1.

Tests were focused on the reliability and sensitivity of the localization tool to errors in the data obtained from possible sensors located in the field. This kind of test was performed to check the sensitivity on the localization method results on the accuracy of the measured data regarding obtained pollutograms. For this reason, we intentionally inserted random and systematic errors into the modelled data in the given range of percent (0%, 5% and 10%). For examples of the generated data with

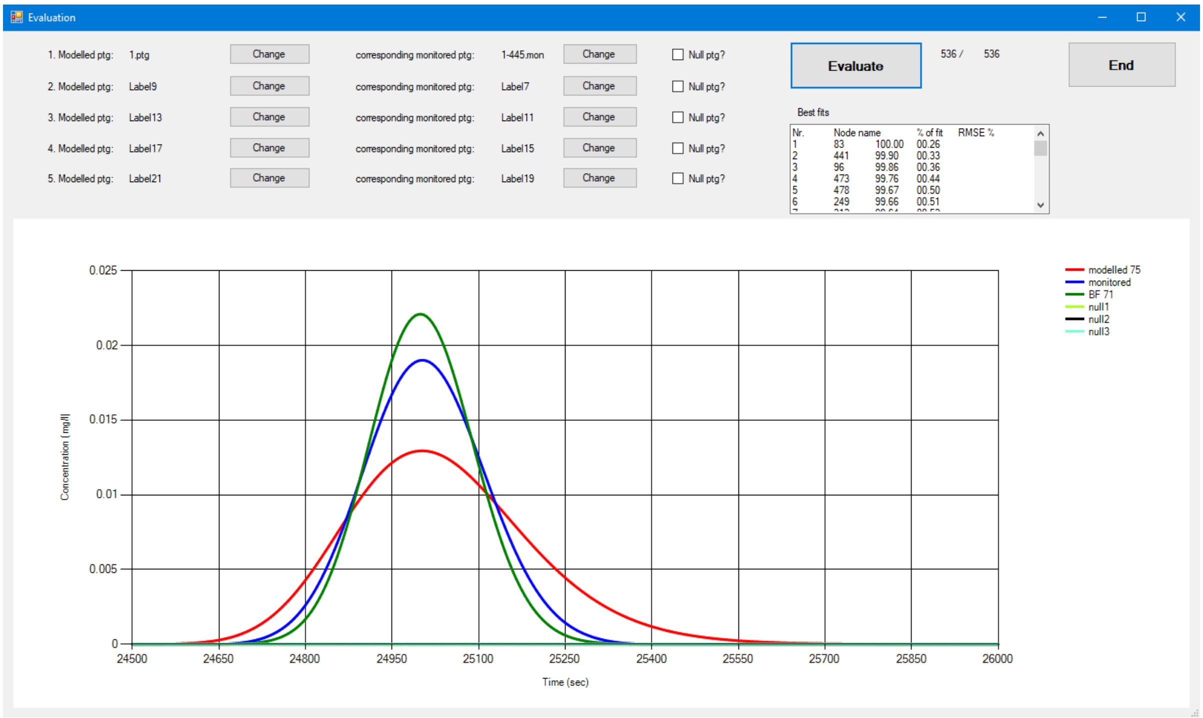

Fig. 2. Screenshot of the localisation SW procedure.

Table 1. Summary of sewer network parameters used for numerical test

Catchment	
Total catchment area	131.51 ha
Total impervious area	31.74 ha
Average runoff coefficient	0.241
Number of elementary sub-catchments	448
Min. node invert level	294.00 m a.s.l.
Max. node invert level	340.05 m a.s.l.
Sewer network	
Number of nodes (manholes)	562 (279*)
Number of sections (branches)	615 (261*)
Total sewer length	26 066 m (11 062*)

random and systematic error see Fig. 3 and Fig. 4. Such modified pollutogram was then used as input into the localization procedure as the basic input (instead of correct generated data). Such errors in the measured data can be caused not by the inaccuracy of the measuring devices, but also by the flow and concentration fluctuations.

Results and discussion

In our case of the sewer network (see Fig. 5), as the monitored manhole/profile was used the manhole Nr. 1 – the inlet manhole at the waste water treatment plant (WWTP). The tests were performed for pollutant inlet in the manhole Nr. 54 (see Fig. 5). The manhole Nr. 54 is 2622 m far from the WWTP (flow route on the sewer network). For the numerical test we used the initial (original) data, achieved by numerical modelling of the pollution spreading in the open channel system, namely for the manhole Nr. 54 (see Fig. 5). For the sensitivity analysis we apply on this dataset various (virtual) types of errors, as described later.

The initial test of the localisation procedure, using with zero error (0%, i.e. with the original data) selected the manhole Nr. 54 as the location with the highest probability of the source location. This was also a confirmation that the methodology is correct and the localization procedure was coded correctly.

However, this simple numeric test also showed other locations that had a high probability of locating the substance source in the interval chosen (90%). These locations in Fig. 5 are shown by the red colour of the points. As can be seen in this figure, there are several such locations on connected (side) open channels. For this reason, it is advisable to establish monitoring points upstream the connection of such channels to the main channel; this can in real cases confirm or exclude the inflow of pollution from a given direction by detection (or absence) of monitored substance.

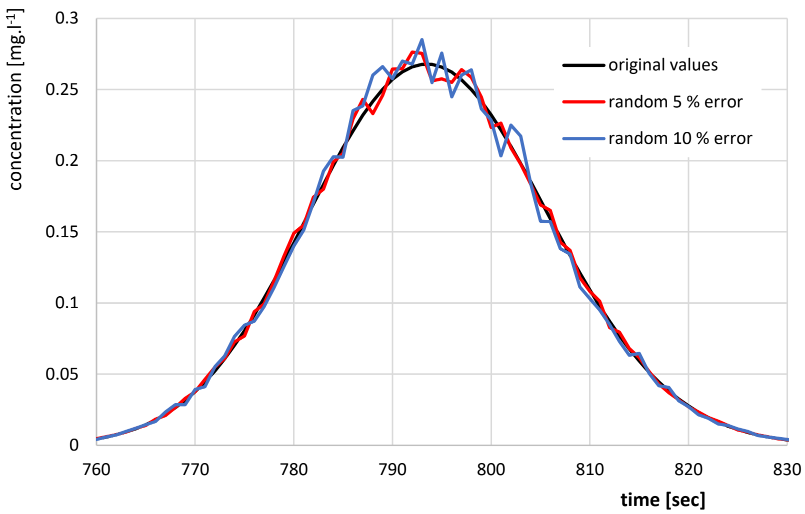

Fig. 3. Example of random error generation.

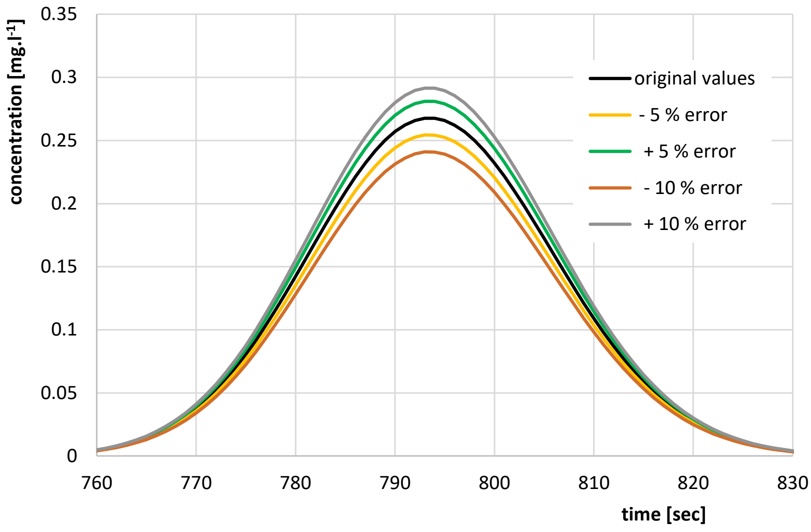

Fig. 4. Example of systematic error generation.

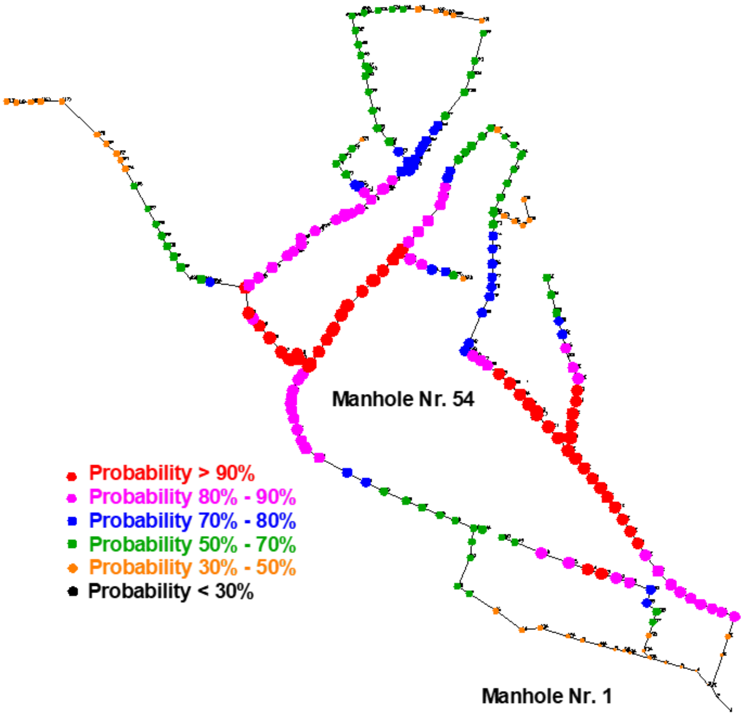

Fig. 5. Sewer network system used for tests, location probability results for zero error.

The tests with random errors (see Fig. 3) show, that small errors in measured values with ±5% does not significantly influence the accuracy of the localisation method. In opposite, larger errors in measured values (±10%) cause significantly higher localisation errors. Typically, the substance source dropped in the probability ranking by 34 up to 56 places in the highest probability ranking list, but the relative probability remains still very high (96.6–99.6%).
Analysing the data, we tried to suppress the random errors down applying the method of the moving average, using different number of the array members for the calculation of the moving average (3, 5 and 10 array members). This does not improve the results accuracy in any way, the accuracy was rather worse.
A second test was also performed, with simulated systematic error, i.e. the modelled data were modified by a systematic error of a certain percentage of the measured value (+5, 10% and –5, 10%, see the Fig. 4). The sensitivity test results that such error has almost no influence on the accuracy of the localisation procedure. Reason of this fact is probably the transformation of the pollutograms to the unitary form, which suppress the errors down.
Of course, the solution or prediction of the pollutant source location also depends on the accuracy of the all input data. There could be another data errors, which can cause large errors in the inverse problem solution, e.g. errors in the open channel physical definitions, their hydraulic data, flow regime (flow rates) etc. For this reason, it is necessary to approach the acquisition of this data with the highest possible degree of accuracy. At the same time, the authors recommend at least to verify the numerical model setup with a small number of tracer experiments in order to verify the parameters of the numerical model and to avoid large errors in the source localization solution.
Comparing our results with the similar studies is not easy, because of lack of relevant papers, dealing with the inverse- source localisation problem in extensive channel systems. The source localisation problem was studied by (Ghane et al., 2016) in the river network, where only few possible source localities were assumed, the study of (Telci and Aral, 2011) assumes 100 possible pollution source points. However, the last one study uses for the pollution point localisation training followed by a sequential elimination of the candidate spill locations which lead to the identification of potential spill locations.
Very similar work to our study was published by (Sonnenwald et al., 2023). The results of this study fully confirmed our assumption, that every point (possible pollution source) in the sewer network has its specific “fingerprint”, which describes the possible concentration distributions observed at a monitoring location as a result of a pulse input at every upstream manhole. This inspiring study also states the fact that it is necessary to take into account the change of hydraulic parameters in the sewer network based on the time of day, or daily activity profile of residents (wastewater production).

Conclusion

In the paper we presented a solution of an inverse problem – the pollution source localisation in a system of open channel system. The proposed inverse problem solution method, as well as the overall localisation procedure was numerically tested on a simplified real sewer system data, whereas part of the numerical tests was focused to the sensitivity analysis of the proposed method.

Numerical tests have confirmed the methodology correctness and the proper coding of the localization procedure. This was confirmed by the initial numeric test, which was performed with the initial (original) data, achieved by numerical modelling of the pollution spreading in the open channel system assuming as the pollution source the manhole Nr. 54. The initial test selects exactly this manhole as the node with the highest probability of the pollution source occurrence.

The sensitivity analysis was performed assuming two error types - random and systematic error. The first one occurs very often in open channels, particularly in sewer networks and is typically caused by probe (sensor) clogging. The sensitivity analysis results show, that this kind of error, especially when the random errors are bigger than 10%, significantly affects the results of the localisation procedure. Use of simply data filters (moving average) does not improve the results, so in this case maybe more advanced techniques shall be used, e.g. the Kalman filter (Chachuła et al. 2021)

The second error type – the systematic error can arise e.g. from incorrect calibration of the measuring device (probe), when the device shows a constant proportionally higher (or lower) value. The sensitivity analysis results of the proposed localisation procedure show, that in this case the input data inaccuracies has almost no influence on the localisation accuracy. The reason for this is the linear transformation of the measured data to the unitary form, which almost eliminates these errors.

Numerical tests confirmed the correctness of the methodology as well as the localization procedure itself. At the same time, however, they pointed out its weak points – random errors in measurements. Besides this type of errors, it is necessary in the next future to deal with other sources of random errors, such as the definitions of the physical properties of the network of open channels, their hydraulic parameters, eventually to perform a physical verification of the proposed model and methodology.

Acknowledgement

This study was supported by the grants from the Slovak Academy of Sciences (project VEGA No. 2/0085/20 with the title "Prediction of a point pollution source position in a watercourse network – a hydrodynamic approach") and by the H2020 EU funded project "SYSTEM", grant agreement No. 787128.

References

Bagtzoglou, A. C., Atmadja, J. (2003): Marching-jury backward beam equation and quasi-reversibility methods for hydrologic inversion: Application to contaminant plume spatial distribution recovery. Water Resour. Res. 39, 1–14. https://doi.org/10.1029/2001WR001021

Banik, B. K., Di Cristo, C., Leopardi, A. (2014): "SWMM5 toolkit development for pollution source identification in sewer systems." Procedia Engineering 89: 750–757. https://doi.org/10.1016/j.proeng.2014.11.503

Banik, B. K., Di Cristo, C., Leopardi, A. (2015): "A pre-screening procedure for pollution source identification in sewer systems." Procedia Engineering 119: 360–369. https://doi.org/10.1016/j.proeng.2015.08.896

Boano, F., Revelli, R. Ridolfi, L. (2005): Source identification in river pollution problems: A geostatistical approach. Water resources research, 41.

Chachuła, Krystian, Robert Nowak, and Fernando Solano. (2021): Pollution Source Localization in Wastewater Networks. Sensors 21, no. 3: 826. https://doi.org/10.3390/s21030826

Courant, R., Hilbert, D. (1993): Methoden der mathematischen Physik, Methoden der mathematischen Physik. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-58039-0

El Badia, A., Hamdi, A. (2007): Inverse source problem in an advection-dispersion-reaction system: Application to water pollution. Inverse Probl. 23, 2103–2120. https://doi.org/10.1088/0266-5611/23/5/017

Fischer, H.B., List, E.J., Koh, R.C.Y., Imberger, J., Brooks, N.H. (1979): Mixing in inland and coastal waters.

Ghane, A., Mazaheri, M., Mohammad Vali Samani, J. (2016): Location and release time identification of pollution point source in river networks based on the Backward Probability Method. J. Environ. Manage. 180, 164–171. https://doi.org/10.1016/j.jenvman.2016.05.015

Hadamard, J. (1923): Lectures on Cauchy's Problem in Linear Partial Differential Equations. New Haven: Yale University Press.

Skaggs, T. H., Kabala Z. J. (1994): Recovering the release history of a groundwater contaminant. Water Resour. Res, 30(1), 71–79.

Cheng, W. P., Jia, Y. (2010): Identification of contaminant point source in surface waters based on backward location probability density function method. Adv. Water Resour. 33, 397–410. https://doi.org/10.1016/j.advwatres.2010.01.004

Kessler, A., Ostfeld, A., Sinai, G., (1998): Detecting accidental contaminations in municipal water networks, J. Water Resour Plan and Manage. 124, 192–198.

Kováčová, V. (2023): Deterioration of water quality in aquatic system. Acta Hydrologica Slovaca, Vol. 24, No. 1, 2023, 141–150, https://doi: 10.31577/ahs-2023-0024.01.0016

Barati Moghaddam, M., Mazaheri, M., Mohammad Vali Samani, J. (2021): Inverse modeling of contaminant transport for pollution source identification in surface and groundwaters: a review. Groundwater for Sustainable Development, 15, 100651. https://doi.org/10.1016/J.GSD.2021.100651

Mazaheri, M., Mohammad Vali Samani, J., Samani, H. M. V. (2015): Mathematical Model for Pollution Source Identification in Rivers. Environ. Forensics 16, 310–321. https://doi.org/10.1080/15275922.2015.1059391

Novák, V. (2022): Global changes and hydrosphere, Acta Hydrologica Slovaca, Vol. 23, No. 1, 2022, 3–9, doi: 10.31577/ahs-2022-0023.01.0001

Runkel, R. L., Broshears, R. E. (1991): One-dimensional transport with inflow and storage (OTIS): A solute transport model for small streams. CADSWES-Center Adv. Decis. Support Water Environ. Syst. - Dep. Civ. Eng. - Univ. Color. 91.

Shao, Z., Xu, L., Chai, H., Yost, S. A., Zheng, Z., Wu, Z., He, Q. (2021): "A Bayesian-SWMM coupled stochastic model developed to reconstruct the complete profile of an unknown discharging incidence in sewer networks." Journal of Environmental Management 297: 113211.

Sokáč, M., Velísková, Y., Gualtieri, C. (2018): An approximate method for 1-D simulation of pollution transport in streams with dead zones. Journal of Hydrology and Hydromechanics 66, no.4 (2018): 437–447. https://doi.org/10.2478/johh-2018-0035

Sonnenwald, F., Shuttleworth, J., Bailey, O., Williams, M., Frankland, J., Rhead, B., Mark, O., Wade, M. J., Guymer, I. (2023): Quantifying Mixing in Sewer Networks for Source Localization. Journal of Environmental Engineering, 149(5): https://doi.org/10.1061/JOEEDU.EEENG-7134

Telci, I. T., Aral, M. M. (2011): Contaminant Source Location Identification in River Networks Using Water Quality Monitoring Systems for Exposure Analysis. Water Qual. Expo. Heal. 2, 205–218. https://doi.org/10.1007/s12403-011-0039-6

Wang, J., Zhao, J., Lei, X., Wang, H. (2018): New approach for point pollution source identification in rivers based on the backward probability method. Environ. Pollut. 241, 759–774. https://doi.org/10.1016/j.envpol.2018.05.093

Woodbury, A., Sudicky, E., Ulrych, T. J., Ludwig, R. (1998): Three-dimensional plume source reconstruction using minimum relative entropy inversion. J. Contam. Hydrol. 32, 131–158. https://doi.org/10.1016/S0169-7722(97)00088-0

Ing. Yvetta Velísková, PhD. (*corresponding author, e-mail: veliskova@uh.savba.sk)
Assoc. Prof. Ing. Marek Sokáč, PhD.
Institute of Hydrology SAS
Dúbravská cesta 9
841 04 Bratislava
Slovak Republic

Maryam Barati Moghaddam, PhD.
Department of Water Engineering and Management
Faculty of Agriculture
Tarbiat Modares University of Tehran
Irán