Runoff prediction using rainfall data from microwave links: Tabor case study
David Stransky, Martin Fencl and Vojtech Bares

ABSTRACT
Rainfall spatio-temporal distribution is of great concern for rainfall-runoff modellers. Standard rainfall observations are, however, often scarce and/or expensive to obtain. Thus, rainfall observations from non-traditional sensors such as commercial microwave links (CMLs) represent a promising alternative. In this paper, rainfall observations from a municipal rain gauge (RG) monitoring network were complemented by CMLs and used as an input to a standard urban drainage model operated by the water utility of the Tabor agglomeration (CZ). Two rainfall datasets were used for runoff predictions: (i) the municipal RG network, i.e. the observation layout used by the water utility, and (ii) CMLs adjusted by the municipal RGs. The performance was evaluated in terms of runoff volumes and hydrograph shapes. The use of CMLs did not lead to distinctly better predictions in terms of runoff volumes; however, CMLs outperformed RGs used alone when reproducing a hydrograph’s dynamics (peak discharges, Nash–Sutcliffe coefficient and hydrograph’s rising limb timing). This finding is promising for number of urban drainage tasks working with dynamics of the flow. Moreover, CML data can be obtained from a telecommunication operator’s data cloud at virtually no cost. That makes their use attractive for cities unable to improve their monitoring infrastructure for economic or organizational reasons.

Key words | commercial microwave links, quantitative precipitation estimates, rainfall monitoring, rainfall-runoff modelling, storm runoff prediction, urban hydrology

INTRODUCTION
Rainfall-runoff models, a basic hydroinformatic tool of urban drainage management, have been used for almost three decades (Vojinovic & Abbott 2017) and are of great help in operational tasks as they enable the optimization of urban drainage systems and more effective planning of their development. However, model performance is affected by uncertainties arising from input data, model structure and identifying model parameters (Renard et al. 2010) which can lead to wrong decisions and to an inefficient drainage operation and unexpected failures as a result. Thus, it is crucial not only to carefully construct and calibrate the rainfall-runoff model but also to input reliable rainfall data. However, this requires dense and well-maintained rainfall observation networks (Schilling 1991).

The density of rain gauges (RGs) in the networks is usually sufficient for reliable runoff prediction only in the case of experimental setups or short-term installations, e.g. Niemczynowicz & Dahlblom (1984) (12 RGs in a 20 km² catchment), Lei (1996) (seven RGs in a 0.4 km² catchment) and Fencl et al. (2017) (six RGs in a 1.24 km² catchment). However, RG density in networks used for operational purposes (i.e. permanent installations) is usually much sparser for practical and financial reasons (Berne et al. 2004) and is often insufficient for runoff predictions, especially during convective storms, which exhibit strong spatial variations over short distances (Goudenhoofdt et al. 2017).

Completing ground observations by weather radar would be a logical option, though it has several major disadvantages: (i) weather radar measures hundreds of metres above ground, (ii) it is affected by a variety of errors, and (iii) it is not often available and new installations are rather expensive (Thorndahl et al. 2017). Therefore, a novel, economically viable source of rainfall data is needed for operational applications (Rabiei et al. 2016). Most studies seek alternatives which are either not initially intended for rainfall estimation or have low operational...
costs, e.g. acoustic rain gauges (de Jong 2010), RainCars (Haberlandt & Sester 2010) and microwave links (Upton et al. 2005). Commercial microwave links (CMLs) in particular have attracted wider interest in the hydrological and meteorological research community in the last decade.

CMLs are radio connections widely used to connect cellular network nodes, thus their density is especially high in urbanized catchments where data traffic is generally high. CMLs operate at millimetre wavelengths where radio waves are substantially attenuated by raindrops (Olsen et al. 1978). This attenuation can be related to rainfall intensity along the link path (Messer et al. 2006):

\[ R = a \cdot k^\beta \]  

(1)

where \( R \) [mm/h] is rainfall intensity, \( k \) [dB/km] specific attenuation of CML caused by raindrops along its path and \( a \) and \( \beta \) empirical parameters dependent on CML frequency, polarization and drop size distribution. The CML rainfall estimates are, however, often biased due to a variety of other factors influencing CML attenuation (Leijnske et al. 2008) which hinders proper identification of \( k \). Different uncertainties affecting CML rainfall estimates have been addressed in specially designed experimental setups (e.g. Fenicia et al. 2012; Schleiss et al. 2013).

Fencl et al. (2017) showed that CML data can be combined with rain gauges to provide reliable high-resolution quantitative precipitation estimates (QPEs) in a pilot catchment in Prague – Letnany, in the Czech Republic. However, it is not clear to what degree the method is suitable for routine modelling tasks and is able to improve rainfall-runoff predictions. Moreover, CML density in the Letnany pilot catchment is an order of magnitude higher (\( \approx 15 \) CMLs/km\(^2\)) than the average CML density in the majority of Czech cities (\( \approx 1.4–2.7 \) CMLs/km\(^2\)) (Fencl et al. 2017). Nevertheless, even sparser CML networks have substantially higher densities than RG networks (typically 0.05–0.10 RGs/km\(^2\) in larger Czech cities) and could therefore conveniently complete existing rainfall monitoring networks.

In this work, we use CML data as a supplement to a municipal monitoring scheme (a permanent monitoring network used for a drainage network operation) and compare simulated and measured runoffs. Only models and data (except CMLs) provided by the water utility are used, i.e. we tested the extent to which CMLs can be integrated into standard non real-time tools used for urban drainage management, and whether they can improve the performance of those tools. The case study was carried out in a catchment of the Tabor agglomeration with a CML density typical of the majority of Czech cities.

**METHODS**

The rainfall-runoff model (MIKE URBAN, MOUSE model, DHI) of the Tabor agglomeration was provided by the water utility. Runoff hydrographs were simulated using rainfall data from (i) the municipal RG network alone and (ii) CMLs adjusted by municipal RG network observations. The reliability of simulations for each rainfall observation layout was evaluated by comparison with the measured runoff.

**Catchment**

An agglomeration of the cities of Tabor, Sezimovo Usti and Plana nad Luznici (a catchment area of 16 km\(^2\), 46,000 inhabitants) was selected for the case study. The catchment has diverse land uses and is drained by a combined sewer system with a total length of 209 km. The drainage system is connected to two wastewater treatment plants (WWTPs) for 20,000 (WWTP1 – Klokoty) and 90,000 PE (WWTP2 – ACOV) respectively.

**Municipal monitoring network**

The municipal urban drainage monitoring network, established in 2013–2014, is extensive in comparison to similar sized or even larger cities around the world. It consists of five permanent tipping bucket RGs and seven flowmeters, supplemented by several portable flow measurements for ad hoc purposes.

In the present study, data from all RGs and four flowmeters covering the largest sewer sub-catchments were used. Flowmeters of interest are located as shown in Figure 1: (FL1) WWTP1 inflow, (FL2) Old Town and Klokoty catchment, (FL3) Horky catchment, (FL4) WWTP2 inflow.

**CML data**

There are 15 CMLs operated by T-Mobile, CZ in the agglomeration (see Figure 2, Table 1). CML attenuation data are collected in near real-time at a sampling frequency of approximately 10 seconds using a specific server-sided software application running at the network operational centre at T-Mobile headquarters in Prague. The application actively polls selected CMLs using simple network management...
protocol (SNMP) commands and stores the records in an SQL database (Fencl et al. 2013). CML attenuation data are processed and adjusted according to Fencl et al. (2013) using data from five permanent RGs aggregated to 15-min intervals resulting in QPEs at 1-min temporal resolution for each CML.

Rainfall-runoff model

The rainfall-runoff was simulated using a detailed rainfall-runoff model. The model was constructed, calibrated and verified within the agglomeration urban drainage master
plan in 2007 and consists of 2,117 nodes, 62 combined sewer overflows and 12 pumping stations.

Two rainfall inputs were used: (i) municipal RG network alone and (ii) CMLs adjusted by RGs. The evaluated rainfall events occurred from spring to autumn 2014. For the performance evaluation, 11 rain events were selected (Table 2). The most intense rainfall event (May 23) has a 20-year return period (30-min duration window) and the rainfall event with highest rainfall depth (May 27) has a more than 10-year return period (240-min duration window).

### Assigning rainfall to the model

The CMLs have different lengths and are distributed unevenly over the catchment. To assign the CML path-averaged rainfall information to the model sub-catchments, the catchment was divided into eight regions representing the sewer system topology and the CML network topology (Figure 2).

The regions close to the city centre (a–d) are well-covered by CMLs. Moreover, each CML crosses only one region. Each CML can, therefore, be directly assigned to the region being crossed by its path. CML ID 6 is excluded as 80% of its path extends beyond the region. The QPEs from CMLs belonging to the same region are averaged as simple means. The remaining regions (e–h) are covered mostly by long CMLs extending outside of their areas and crossing several regions or (in the case of region e) not covered at all. To assign the CMLs to these regions reliably, the following assumptions were made: (i) the most representative CML for an uncovered catchment was the closest CML with a length scale similar to the region, (ii) the CML crossing

### Table 1 | Characteristics of CMLs in the Tabor agglomeration catchment

<table>
<thead>
<tr>
<th>ID</th>
<th>Frequency (GHz)</th>
<th>Azimuth</th>
<th>Length (m)</th>
<th>ID</th>
<th>Frequency (GHz)</th>
<th>Azimuth</th>
<th>Length (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>279°</td>
<td>601</td>
<td>9</td>
<td>38</td>
<td>107°</td>
<td>957</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>310°</td>
<td>1,365</td>
<td>10</td>
<td>32</td>
<td>118°</td>
<td>1,673</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>68°</td>
<td>1,363</td>
<td>11</td>
<td>38</td>
<td>124°</td>
<td>1,471</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>81°</td>
<td>5,274</td>
<td>12</td>
<td>25</td>
<td>150°</td>
<td>5,078</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>95°</td>
<td>514</td>
<td>13</td>
<td>38</td>
<td>185°</td>
<td>266</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>96°</td>
<td>10,257</td>
<td>14</td>
<td>25</td>
<td>153°</td>
<td>2,742</td>
</tr>
<tr>
<td>7</td>
<td>38</td>
<td>104°</td>
<td>1,950</td>
<td>15</td>
<td>23</td>
<td>157°</td>
<td>7,269</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>110°</td>
<td>2,075</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The origin of azimuth vectors was set to the CML hub in the city centre with the exception of CML ID 8 (vector origin was placed about 230 m east of the hub as the antenna is mounted on a different building than the hub) and CML ID 13 (short link in Sezimovo Usti, vector origin was set to northern node of the link). Polarization of CMLs is unknown, but were not of interest as QPEs were adjusted by ground measurements provided by municipal RG network.

### Table 2 | Rain events used for the performance evaluation

<table>
<thead>
<tr>
<th>No.</th>
<th>Date</th>
<th>Duration (min)</th>
<th>Depth (mm)</th>
<th>Max. 1-min intensity (mm/h)</th>
<th>Average intensity (mm/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>May 23</td>
<td>0–45</td>
<td>0–28.8</td>
<td>0–180</td>
<td>0–71</td>
</tr>
<tr>
<td>2</td>
<td>May 27</td>
<td>180–260</td>
<td>17.5–64.4</td>
<td>18–204</td>
<td>3.5–14.8</td>
</tr>
<tr>
<td>3</td>
<td>Jul 21</td>
<td>254–350</td>
<td>8.3–25.2</td>
<td>30–78</td>
<td>1.6–4.1</td>
</tr>
<tr>
<td>4</td>
<td>Jul 28</td>
<td>17–37</td>
<td>5.6–11.5</td>
<td>42–84</td>
<td>11.8–31.3</td>
</tr>
<tr>
<td>5</td>
<td>Jul 30</td>
<td>81–149</td>
<td>5.5–26.5</td>
<td>30–66</td>
<td>2.9–18.9</td>
</tr>
<tr>
<td>6</td>
<td>Aug 04</td>
<td>26–46</td>
<td>5.4–26.7</td>
<td>36–192</td>
<td>11.2–35.9</td>
</tr>
<tr>
<td>7</td>
<td>Aug 11</td>
<td>418–464</td>
<td>9.7–10.6</td>
<td>6–36</td>
<td>1.3–1.5</td>
</tr>
<tr>
<td>8</td>
<td>Aug 27</td>
<td>593–789</td>
<td>20.5–25.3</td>
<td>12–36</td>
<td>1.7–2.3</td>
</tr>
<tr>
<td>9</td>
<td>Aug 31</td>
<td>188–578</td>
<td>16.0–24.9</td>
<td>12–48</td>
<td>2.4–8.0</td>
</tr>
<tr>
<td>10</td>
<td>Sep 01</td>
<td>1298–1708</td>
<td>12.3–15.2</td>
<td>6–12</td>
<td>0.5–0.6</td>
</tr>
<tr>
<td>11</td>
<td>Sep 20</td>
<td>54–451</td>
<td>2.5–7.7</td>
<td>6–42</td>
<td>0.8–2.8</td>
</tr>
</tbody>
</table>

The range of values describes the variability between municipal network rain gauges. All rainfall events occurred in 2014.
more regions was most representative for a region close to its midpoint, and (iii) long CML were divided into virtual sections \( s_1 \) and \( s_2 \) and rainfall distribution along these sections was estimated using observations from a nearby CML (Goldshstein et al. 2009). In our case, it was convenient to divide CML ID 15 into two sections. The first section \( s_1 \) corresponds to the length of the shorter CML ID 12 with a joint node and similar azimuth but shorter path, while the second section \( s_2 \) corresponds to the rest of the CML ID 15 path (Figure 3), thus representing region g relatively well. The rainfall intensity along section \( s_1 \) was considered to be the same as the rainfall intensity along the shorter CML ID 12. Therefore, the rainfall intensity of section \( s_2 \) can be expressed and calculated from Equations (2)–(4).

\[
R_{CML ID 15} = \frac{R_{CML ID 15(s_1)} \cdot L_{CML ID 15 (s_1)} + R_{CML ID 15(s_2)} \cdot L_{CML ID 15 (s_2)}}{L_{CML ID 15}}
\]

(2)

\[
R_{CML ID 15 (s_1)} = R_{CML ID 12}
\]

(3)

\[
L_{CML ID 15 (s_1)} = L_{CML ID 12}
\]

(4)

where \( R \) is the rainfall intensity along the CML path (or its section) and \( L \) is the length of a CML (or its section). CML data assignment to the regions is summarized in Table 3.

The municipal RG network data are assigned to the model sub-catchments using Thiessen polygons (Figure 1).

**Performance criteria**

The agreement of each simulated runoff hydrograph with the measured one was evaluated using following performance criteria: (i) relative error between the hydrograph volumes \( (\Delta V) \), (ii) relative error between hydrograph peak discharges \( (\Delta Q_{max}) \), (iii) Nash-Sutcliffe model efficiency coefficient \( (\text{NSE}) \), (iv) time shift between the start of the discharge rises \( (\Delta T_{(s)}) \), and (v) time shift between the end of the discharge rises \( (\Delta T_{(e)}) \) (Figure 4). Peak discharge statistics were calculated for events with at least 15% pipe filling to avoid large relative errors during low flows.

The start of the discharge rise was defined as the time of the first discharge value higher than the dry-weather base flow, followed by a further increase in the value of the discharge. The end of the discharge rise was defined either as the time of peak discharge value for sharp-peaked hydrographs, or as the time when a gradient of a hydrograph's

### Table 3 | CML data assignment to regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Relevant CML ID</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>3, 4</td>
<td>Rainfall intensity averaged from relevant links</td>
</tr>
<tr>
<td>b</td>
<td>5, 7, 8, 9, 10, 11</td>
<td>Rainfall intensity averaged from relevant links</td>
</tr>
<tr>
<td>c</td>
<td>–</td>
<td>Region does not affect flow in monitored locations</td>
</tr>
<tr>
<td>d</td>
<td>1, 2</td>
<td>Rainfall intensity averaged from relevant links</td>
</tr>
<tr>
<td>e</td>
<td>14</td>
<td>CML lies outside the region, but is closest of all CMLs</td>
</tr>
<tr>
<td>f</td>
<td>12, 13</td>
<td>Rainfall intensity averaged from relevant links</td>
</tr>
<tr>
<td>g</td>
<td>–</td>
<td>Rainfall intensity averaged from regions f and h</td>
</tr>
<tr>
<td>h</td>
<td>15, (12)</td>
<td>CML 15 divided into two sections with help of CML ID 12, see Equations (2)–(4)</td>
</tr>
</tbody>
</table>

CML ID 6 was not used as only 20% of its length lies within the urbanized catchment.
rising limb decreased to zero for ‘flat-top’ hydrographs. Nevertheless, manual specification was needed in several cases of ‘flat-top’ hydrographs, where slight inaccuracy in the gradient of simulated runoff might cause gross error in the time shift although the overall characteristics of the hydrograph were very well reproduced.

In total, 44 pairs of hydrographs were investigated, i.e. 11 rain events in four discharge monitoring locations. In 11 cases, measured discharge data were not representative (e.g. gaps in measurement and high data noise). Therefore, the evaluation was performed for 33 hydrographs.

**RESULTS AND DISCUSSION**

No distinctive difference between hydrograph volumes simulated using the municipal RG network and CMLs was observed as the CML attenuation rainfall model was adjusted to the RG observations (Table 4). The peak discharges are, however, better reproduced by rainfall observations complemented by CMLs with \( \Delta Q_{\text{max}} = 1.7\% \) (\( \sigma = 11.4\% \)) compared to \( \Delta Q_{\text{max}} = 8.0\% \) (\( \sigma = 19.4\% \)) when RG observations were used alone (Table 4). Similarly, use of CMLs leads to substantially better performance in terms of NSE. The substantial increase of NSE was observed at the monitoring points FL1 and FL2 (Table 4).

Further, completing municipal RG network with CMLs substantially improved the timing of the hydrograph’s rising limb of the simulated runoff (Figure 5). 52% of hydrographs showed better timing for CML-based simulations, 36% of hydrographs showed no distinctive differences between both datasets. Finally, by 12% of hydrographs only the rising limb was better reproduced by RGs alone. Results for all 33 studied hydrographs are shown in a scatter graph in Figure 6.

Results based on a municipal RG network showed an earlier rise of discharge as most events were in the lower-left quadrant of the scatter graph. This bias was site-specific as it depended on terrain morphology, the topology of the permanent RG network and prevailing meteorological

---

**Table 4** | Average difference in volume, peak discharge and average Nash–Sutcliffe model efficiency coefficient for municipal RG-based hydrographs and CML-based hydrographs (compared to monitoring data for 11 rainfall events)

<table>
<thead>
<tr>
<th>Monitoring point</th>
<th>FL1</th>
<th>FL2</th>
<th>FL3</th>
<th>FL4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta V_{\text{RG}} ) [%]</td>
<td>+6.4</td>
<td>+4.3</td>
<td>+9.2</td>
<td>+6.2</td>
</tr>
<tr>
<td>( \Delta V_{\text{CML}} ) [%]</td>
<td>+6.9</td>
<td>+2.9</td>
<td>+12.0</td>
<td>+8.2</td>
</tr>
<tr>
<td>( \Delta Q_{\text{max,RG}} ) [%]</td>
<td>+10.3</td>
<td>-a</td>
<td>+7.5</td>
<td>-3.2</td>
</tr>
<tr>
<td>( \Delta Q_{\text{max,CML}} ) [%]</td>
<td>+6.2</td>
<td>-a</td>
<td>+3.3</td>
<td>-7.2</td>
</tr>
<tr>
<td>NSE(_{\text{RG}}) [-]</td>
<td>0.616</td>
<td>0.556</td>
<td>0.507</td>
<td>0.553</td>
</tr>
<tr>
<td>NSE(_{\text{CML}}) [-]</td>
<td>0.785</td>
<td>0.736</td>
<td>0.549</td>
<td>0.569</td>
</tr>
</tbody>
</table>

The small systematic volume overestimation was caused by the original rainfall-runoff model calibration that was made on the ‘safe side’ (i.e. not to underestimate runoff).

*aWhen runoff events with less than 15% pipe filling were excluded, the statistical set for monitoring point FL2 was not representative due to the small number of elements in the set."
conditions. Results based on CMLs were more equally distributed around zero. Furthermore, confidence intervals were much narrower in the case of CML results than in the case of the municipal RG network. This indicates that CMLs reproduce temporal dynamics of rainfall better than RGs, which is consistent with our previous findings (Fencl et al. 2013) as well as previous studies (e.g. Ochoa-Rodriguez et al. 2013), from which it can be concluded that temporal resolution of rainfall information has a greater effect upon hydrodynamic modelling than spatial resolution.

Further, CML-based simulations captured the runoff dynamics at the Horky catchment surprisingly well (Figure 7). This catchment is a specific (region e in Figure 2) as no CML crosses it and the outlying CML ID 14 had to be assigned to the catchment (Figure 2, Table 3). The good performance of CML-based runoff predictions may have been due to the path-integrated CML rainfall intensity representing a larger area than the rainfall intensity measured at one point by a single RG.

Although no extreme rainfall events were monitored during the experimental period, we expect CMLs to be less prone than RG networks to completely missing convective cells as CMLs cover urban catchments better than RGs. Nevertheless, the adjusting algorithm would need to be further developed to also perform satisfactorily for extreme convective rainfall events with high spatial variability.
CML implementation into the monitoring scheme led to the following:

Catching extreme rainfall intensities also depends on CML functional availability which is related to signal attenuation by rain drops. As the requirement for availability is typically set to 99.999% of the time, CMLs can be inoperative up to 5.26 min per year. However, our experience with data collected in 2012–2018 shows that the availability is higher than required.

The above-stated findings show an improvement of runoff prediction by integrating the CML data into the urban drainage monitoring network. Although our results are specific to Tabor’s catchment, we are confident that CMLs can improve the performance of rainfall-runoff models in other cities. Tabor’s municipal urban drainage monitoring network is one of the densest in Czech cities of similar size and corresponds better to the monitoring networks of cities of about 100,000 inhabitants. The effect of additional rainfall information from CMLs in sparsely gauged urban catchments is likely to be even more significant.

There are several options for combining the CML data with the permanent RG network data. One of these options is to use the RG data not only for CML corrections, but also as additional information for rainfall spatial distribution. However, since we used this case study for CML data validation with the help of hydrological modelling, we evaluated the data sources separately to be consistent with our previous work (e.g. Fencl et al. 2017).

CONCLUSIONS

In this paper, we have presented the rainfall-runoff modelling results from a case study in the Tabor agglomeration, Czech Republic, in which the municipal RG network was complemented by CML ‘rainfall sensors’. We concluded the following:

- CML implementation into the monitoring scheme led to improved prediction in the catchment with an above-standard municipal monitoring network and moderate CML coverage.
- CML ‘rainfall sensors’ improved the prediction of hydrograph dynamics in all related performance criteria (peak discharge \(Q_{\text{max}}\), NSE and rising limb timing). This is important for a wide range of urban drainage tasks.
- The results of the presented case study are encouraging, especially because CML data are technically easy to obtain from a telecommunication operator’s data cloud at virtually no cost. Furthermore, the use of the existing CML infrastructure minimizes the risk of physical interference which might occur when extending the rainfall monitoring network with new RGs or local weather radar.

Thus, CML data represent a promising alternative for cities unable to improve their monitoring infrastructure for economic or organizational reasons, and can contribute to improving their urban stormwater management.

ACKNOWLEDGEMENTS

This work was supported by the project of the Czech Science Foundation No. 17-16389S. Special thanks to Vodarenska spolecnost Taborsko and CEVAK for providing us with the rainfall-runoff model and monitoring data and T-Mobile Czech Republic for providing us with CML signal attenuation data.

REFERENCES


Goldstein, O., Messer, H. & Zinevich, A. 2009 Rain rate estimation using measurements from commercial


First received 19 November 2017; accepted in revised form 21 March 2018. Available online 5 April 2018