https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

Hydrological response of a peri-urban catchment exploiting conventional and unconventional rainfall observations: the case study of Lambro catchment

Greta Cazzaniga[1], Carlo De Michele[1], Michele D'Amico[2], Cristina Deidda[1], Antonio Ghezzi[1], and Roberto Nebuloni[3]

[1]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Milan, Italy
[2]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy
[3]Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, Consiglio Nazionale delle Ricerche, Milan, Italy

Correspondence: Carlo De Michele (carlo.demichele@polimi.it), Greta Cazzaniga (greta.cazzaniga@polimi.it)

Abstract. Commercial Microwave Links (CMLs) can be used as opportunistic and unconventional rainfall sensors by converting the received signal level into path-averaged rainfall intensity. Since in meteorology and hydrology the reliable reconstruction of the rainfall spatial distribution is still a challenging issue, there is a wide-spread interest in integrating the precipitation estimates gathered by the ubiquitous CMLs with the conventional rainfall sensors, i.e. rain gauges (RGs) and weather radars. Here we investigate the potential of a dense CML network, for the estimation of river discharges via a semi-distributed hydrological model. The analysis is conducted on Lambro, a peri-urban catchment located in northern Italy and covered by 50 links. A two-level comparison is made between CML- and RG-based outcomes, relying on 12 storm/flood events. First, rainfall data are spatially interpolated and assessed in a set of significant points of the catchment area. Rainfall depth values obtained from CMLs are definitively comparable with direct RG measurements, except for the spells of persistent light rain, due to limited sensitivity of CMLs caused by the coarse quantization step of raw power data. Moreover, it is showed that, when changing the type of rainfall input, a new calibration of model parameters is required. In fact, after the re-calibration of model parameters, CML-driven outputs performances are comparable with RG-driven ones, confirming that the exploitation of a CML network may lead to benefit in hydrological modelling.

1 Introduction

Precipitation is the main downward forcing of the water cycle (Kidd and Huffman, 2011) and consequently one of the most relevant inputs in hydrological models, which are key tools in early warning systems for flood risk forecasting and mitigation (EU Water Directors, 2003). However, precipitation exhibits a significant temporal and spatial variation over a catchment area or region (Dawdy and Bergmann, 1969; Bengtsson, 2011; Parkes et al., 2013) and this is a critical aspect leading to difficulties in reconstructing a reliable rainfall field. In the past, several studies investigated the effects of spatio-temporal variability of rainfall on the hydrological model outputs (e.g., Obled et al., 1994; Bárdossy and Das, 2008; Younger et al., 2009; Arnaud et al., 2011) proving that precipitation inputs have a marked influence on the simulated outflow hydrographs. It is also known

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

that the reconstruction of rainfall input is more accurate as the number of rainfall measurements increases over a study area (Chen et al., 2010; Xu et al., 2013), even if there are some economic or geographical circumstances that prevent the presence of an adequate density of rainfall sensors.

Currently, the most common ground-based technology for rainfall measurement is the rain gauge (RG), which provides single-point measurements (New et al., 2001). In addition, high-precision ground sensors, namely the disdrometers, provide size and velocity of hydrometeors (Jaffrain et al., 2011; Cugerone and De Michele, 2015). One of the major problems encountered when dealing with single-point measurements is to transfer the information in ungauged sites or to reconstruct the rainfall field over the catchment of interest. Such estimates can be performed by the use of spatial interpolation techniques. Several methods are now available, with different degrees of complexity. They can be either deterministic (e.g., the inverse distance weighting (IDW) method (Shepard, 1968) and the Thiessen polygon method (Thiessen, 1911)) or stochastic (e.g., the Kriging technique (Delhomme, 1978) and co-Kriging (Myers, 1984)). However, the outcome of these techniques is proved to be highly sensitive to the gauge density (Xie et al., 1996). Alternatively, the rainfall field at ground can be indirectly obtained by weather radars, when available. The radar retrieves the average rainfall intensity across a volume from measurements of reflectivity through power-law formulas as the one proposed by Marshall and Palmer (1948), see also Raghavan (2013) for a review of reflectivity-rainfall intensity formulas. Ignaccolo and De Michele (2020) and Jameson and Kostinski (2002), have argued about the purely statistical nature of the reflectivity-rainfall intensity formulas, with important consequences about the use of these formulas where calibration with local data is missing. There are in addition other drawbacks associated to the use of radar reflectivity, including the problem of spurious echoes, as ground clutter (Alberoni et al., 2001; Rauber and Nesbitt, 2018), which restrict the use of radars to plain areas and the fact that the radar reflectivity provides only information about precipitable water.

Indeed, measuring the spatial distribution of rainfall is still an open issue, which may be tackled through the integration of the conventional sensors, and/or the complement of new instruments. In this context, the use of opportunistic rainfall sensors is rising a great interest nowadays. Among these, there are the Commercial Microwave Links (CMLs), that are the point-to-point radio links connecting the base stations of a mobile network to the core infrastructure. The use of a mesh of microwave links as rainfall detectors was first proposed by Giuli et al. (1991) through simulations; their method exploits the relationship between the rainfall intensity and the attenuation (i.e., the loss of power) experienced by the electromagnetic wave along the propagation path from the transmitter to the receiver. A pioneering experimental campaign was carried out during the Mantissa project (Rahimi et al., 2003). However, at that time, the need to install ad-hoc microwave links made the technique impractical. A few years later, the scenario changed following the boom in cellular telephony. The use of the ubiquitous CMLs connecting cellular network base stations was first proposed by Messer et al. (2006). Their paper triggered many studies that were conducted worldwide to investigate the CMLs' potential for meteorological and hydrological applications. From a hydrological point of view, CML-based rainfall products were firstly exploited by Fencl et al. (2013) to improve urban drainage modelling in a small scale (2.33 km^2) impervious catchment, in Prague (Czech Republic). Later, Brauer et al. (2016) investigated the effects of the use of CML data in discharge simulations, for a natural low land catchment in the Netherlands, at small scale (6.5 km^2). A further study by Smiatek et al. (2017) used microwave links derived precipitation estimates as rainfall input in a

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

distributed hydrological model applied to the Ammer basin (Germany), at a larger scale (609 km^2). In this work, authors employed the IDW method for interpolation of RG and CML rainfall data on a 100 $\times$ 100 m grid. Another case study was carried out on an agglomeration of cities ($16\,\mathrm{km}^2$) in Czech Republic by Stransky et al. (2018), to check the potential of a dense CML network for urban drainage management. Pastorek et al. (2019) also assessed the impact of both CML sensitivity to rainfall and CML position on urban drainage modelling. Authors found out that the bias propagated thoughout simulations is inversely proportional to CML length. In Italy, Roversi et al. (2020) conducted a meteorological analysis in the Po valley in Emilia Romagna (northern Italy) to assess the accuracy of CML retrieved rainfall rates using data purchased from Vodafone. However, still no one carried out an hydrological application of CML-based rainfall estimates.

Here, the analysis aims at investigating and validating the operational potential of a CML network in a semi-distributed hydrological model in Lombardia (northern Italy). Specifically, we figured out whether rainfall data from a large CML network made up of 50 links may be exploited to provide a reliable reconstruction of the hydrological process and if it is comparable with those achieved with a RG network. To this aim, we investigated a set of summer and autumn precipitation events (both convective and stratiform), that occurred over the Lambro catchment during the years 2019 and 2020. The analysis of events taking place in different seasons allowed us to point out some limits of CMLs in detecting specific types of precipitation. In this work we firstly focused on the spatial interpolation of rainfall observations comparing results from conventional (RGs) and unconventional (CMLs) instruments and their combined use. In fact, the issue of spatial interpolation is crucial when dealing with point (RGs) or linear measurements (CMLs) used as input into a semi-distributed hydrological models, especially when the study area is quite large (in the order of 100 km^2 or even larger). Differently from Stransky et al. (2018), we modified the traditional IDW method to account for CML length in the spatial interpolation. Specifically, in addition to the inverse distance weight, we defined a further weight, which exponentially decreases with CML length. Secondly, we implemented a semi-distributed rainfall-runoff model using three types of inputs: (1) RG measurements, (2) CML estimates and (3) the combination of RG and CML data. We compared results in terms of river discharge.

The remainder of this paper is structured as follows. In Sec. 2 we present the case study, the features of conventional and unconventional sensors networks and the experimental setup. Section 3 includes a description of all methods implemented for the analysis and Sec. 4 reports results including a comparison of rainfall spatial interpolation carried out with the different data types and of stream flow simulations against hydrometric measurements. Discussion and conclusions are given in Sec. 5 and 6, respectively.

2 Case study and experimental setup

The case study is the Lambro catchment, a peri-urban catchment, left side tributary of Po river, in Fig. 1.a. It is located north of Milan metropolitan area and covers three different provinces: Como, Lecco, and Monza and Brianza. Lambro river, at Lesmo river section (in purple in Fig. 1.a), drains an area of 260 km^2 which can be mainly divided in two zones, with different morphology and land use. The northern one is the Prealpine region, where Lambro river rises, at 944 m a.s.l.. The southern one, between Pusiano lake and the outlet section, at 178 m a.s.l., is instead a flat area subjected to massive urbanization, which results

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

in large impervious surfaces and, consequently, fast runoff processes with a lag time of few hours. According to Köppen (1925) climate classification the inland northern portion of Italy belongs to the humid subtropical climate (Cfa). Heavy convective cells characterize the basin, while the highest monthly rainfall accumulation occurs in spring and autumn. The local meteorological drivers, added to urban sprawl, lead to the hydrological vulnerability of the region. In order to mitigate hydrological risk in Monza and Milan urban areas (downstream to our case study), structural works have been carried out along Lambro river in past years. Moreover, great efforts have been put in the implementation and development of non-structural measures providing support to early warning system (e.g., Ravazzani et al., 2016; Masseroni et al., 2017; Lombardi et al., 2018), including a dense monitoring system managed by ARPA Lombardia (Regional Agency for Environmental Protection). From this, we exploited 10 min resolution rainfall depths and temperatures respectively from 13 tipping-bucket RGs and eight thermometers (TH), for years 2018, 2019, and for the first six months of year 2020. In addition, we used 10 min resolution water level measurements of a flow gauge (FG), located at the outlet section of the Lambro basin in the municipality of Lesmo. All these meteorological and hydrological data are available at https://www.arpalombardia.it. A rather dense CML network, owned by Vodafone Italia S.p.A., covers the central and southern catchment area and its surroundings. In contrast, the northernmost portion of Lambro basin is covered by few and unevenly distributed CMLs, given that it is thinly populated and characterized by higher altitudes. The CMLs available over the area are 50. The key features of CMLs as rainfall sensors are the operation frequency and the path length. Regrouping the CMLs according to the frequency:

1. 5 links are in the frequency range [11.4,13.1] GHz, with length between 3.5 and 8 km;
2. 37 links are in the frequency range [18.8,23.0] GHz, with length between 1 and 8.5 km;
3. 8 links are in the frequency range [38.5,42.6] GHz, with length between 1.4 and 2.2 km.

As showed in Fig. 1.a, the catchment area is divided into 15 sub-basins, hereinafter referred to as hydrological response units (HRU), for the sake of implementation of the hydrological model (see Sec. 3.2). In particular, the semi-distributed model here adopted requires, as input data, rainfall depths estimated in the HRU centroids. The estimates were gathered through the IDW technique (see Sec. 3.3) and for each HRU a different number of sensors was exploited. Figure 2 shows some features of the rainfall sensors used for spatial interpolation in each HRU: the number of exploited RGs, CMLs, and their sum, the ratio between CMLs and RGs number, the mean distance between rainfall sensors and HRU centroids, and the mean length of CMLs. It is worth to mention that the total number of available CMLs can be less than the nominal 50 due to maintenance or malfunctioning during some events. The numbers in Fig. 2 are hence averaged over all the events. Figure 2.a shows a significant increase of exploited CMLs from HRU 1 to 6. This could be a potential problem leading to more inaccurate estimates, at the stage of spatial interpolation, for the northern HRUs with respect to the southern ones. On the other hand, the number of RGs has minor variations from one HRU to another. In Fig. 2.b we can see that the lowest ratios between CMLs and RGs correspond to HRU from 1 to 5 and the HRU 10. Moreover, Fig. 2.c shows that, considering HRUs from 1 to 9, the mean distance between sensors and HRU centroids is always higher when CMLs are considered. The opposite trend, with a single exception for HRU 12, occurs for HRUs located further downstream. Lastly, the mean CML length, in Fig. 2.d, has a decreasing trend from upstream to downstream.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

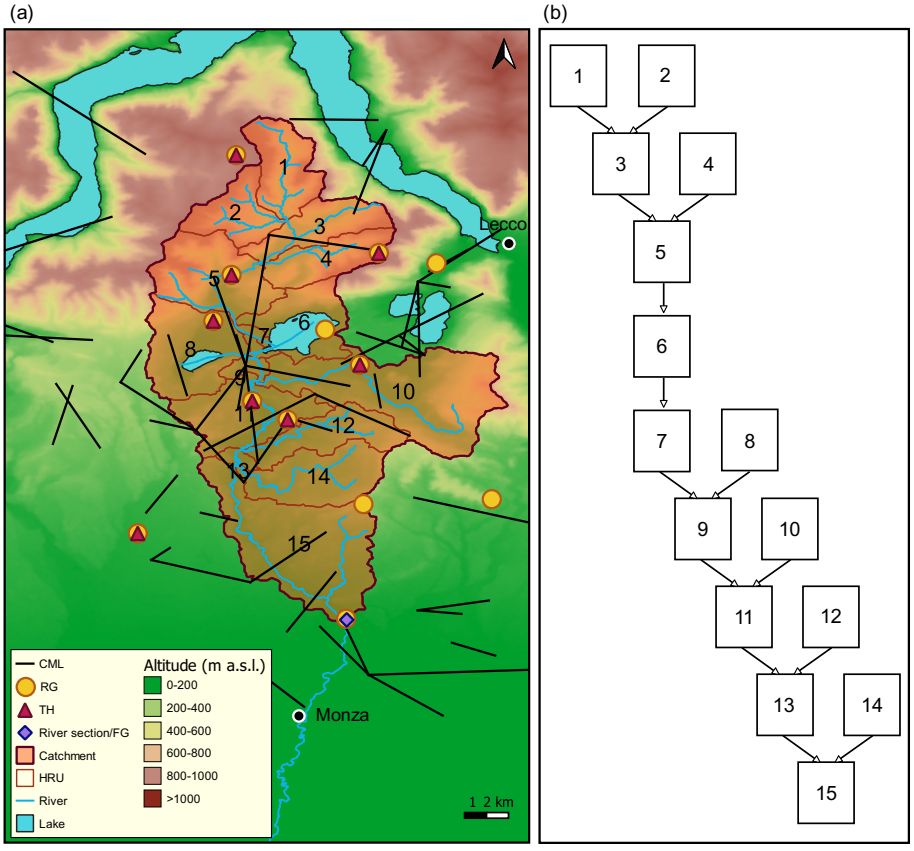

Figure 1. Case study area. Panel (a) shows the Lambro catchment, the partitioning in 15 sub-basins (HRUs), and the position of the sensors, while panel (b) reports the scheme of HRU interaction.

We investigated 12 storm events and the associated floods, in the period June 2019–June 2020. The RG- and CML-based precipitation datasets, aggregated at hourly time scale, cover a wide range of rainy events from summer thunderstorms to low-intensity autumn events. In Table 1 we reported initial and final date and time, accumulated precipitation averaged over 13 RG measurements, maximum detected rain rate and observed total flow volume, for the 12 selected events. We defined a storm event as the time lapse where at least one RG, available on the area, detected precipitation with possible dry intervals no longer than 5 h. A hour is considered dry when the detected rainfall depth is lower than 1 mm and wet otherwise. The beginning of the flood event is conventionally set at the hour in which the flow rate experienced a sudden deviation from the average. The end is instead set when the flow rate reverts to the initial condition, at the end of the depletion curve. According to the maximum observed rain rates, we classified events in *Low rain rate* and *High rain rate*. The former group includes storm events 5, 6, 7, and 12, for which the maximum rain rate is lower than 15 $\mathrm{mm\,h^{-1}}$, while the latter covers the remaining events with maximum rain rate higher than 20 $\mathrm{mm\,h^{-1}}$.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Table 1. Details about the 12 events considered. On the left side, date, time, cumulative precipitation, and maximum rain rate for the 12 storm events. On the right side, date, time, and flow volume of the corresponding flood events.

	Storm event			**Flood event**	
ID event	Date and Time (LT)	Cumulative precipitation (mm)	Max rain rate ($\mathrm{mm\,h^{-1}}$)	Date and Time (LT)	Flow volume ($\mathrm{m^3\ 10^6}$)
1	22 Jun 2019, 06:00 22 Jun 2019, 15:00	40.5	37.6	22 Jun 2019, 08:00 22 Jun 2019, 23:00	0.5
2	14 Jul 2019, 22:00 16 Jul 2019, 03:00	63.6	49.0	22 Jun 2019, 08:00 16 Jul 2019, 23:00	3.0
3	05 Sep 2019, 01:00 09 Sep 2019, 10:00	68.7	36.6	05 Sep 2019, 21:00 11 Sep 2019, 23:00	2.2
4	18 Oct 2019, 17:00 22 Oct 2019, 12:00	108.4	35.2	19 Oct 2019, 00:00 24 Oct 2019, 23:00	8.7
5	14 Nov 2019, 19:00 16 Nov 2019, 17:00	34.5	12.6	15 Nov 2019, 04:00 17 Nov 2019, 04:00	2.9
6	17 Nov 2019, 01:00 17 Nov 2019, 19:00	25.6	6.4	17 Nov 2019, 10:00 19 Nov 2019, 03:00	3.3
7	18 Nov 2019, 23:00 20 Nov 2019, 00:00	27.9	4.0	19 Nov 2019, 04:00 21 Nov 2019, 23:00	5.4
8	14 May 2020, 20:00 16 May 2020, 07:00	64.2	31.0	14 May 2020, 22:00 17 May 2020, 02:00	4.6
9	03 Jun 2020, 16:00 05 Jun 2020, 04:00	79.9	24.8	03 Jun 2020, 18:00 05 Jun 2020, 19:00	4.1
10	07 Jun 2020, 08:00 08 Jun 2020, 02:00	42.1	29	07 Jun 2020, 15:00 08 Jun 2020, 19:00	3.5
11	08 Jun 2020, 17:00 09 Jun 2020, 19:00	32.2	22.8	08 Jun 2020, 20:00 10 Jun 2020, 01:00	3.8
12	10 Jun 2020, 11:00 11 Jun 2020, 07:00	16.5	7.4	11 Jun 2020, 02:00 11 Jun 2020, 15:00	1.0

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

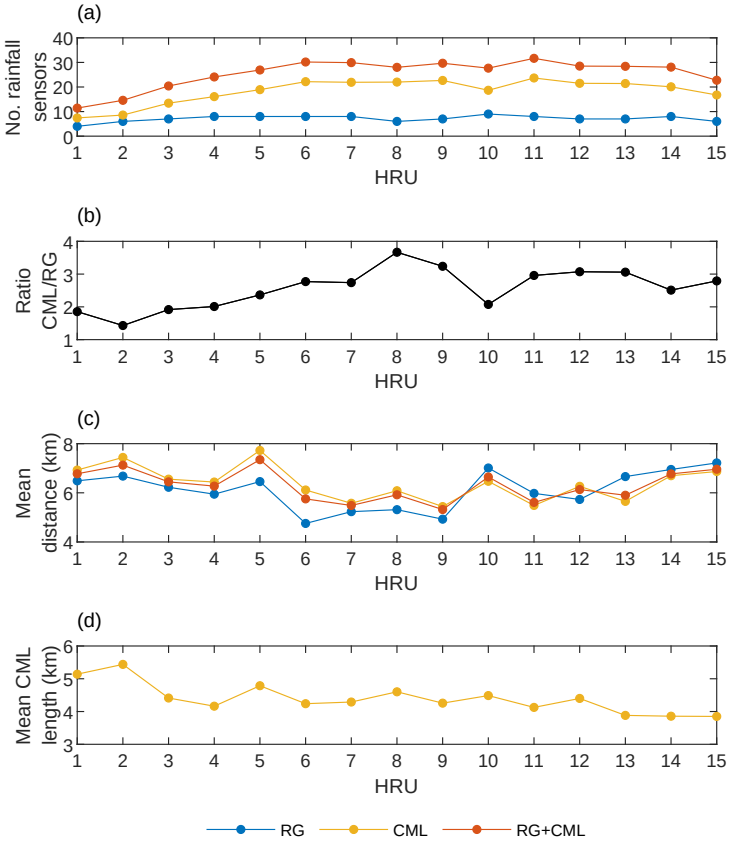

Figure 2. From panel (a) to (d): number of rainfall sensors used for spatial interpolation in each HRU centroid, ratio between the number of RGs and CMLs, mean distance between rainfall sensors and HRU centroids, and CML mean lengths.

3 Methods

In this section we firstly present the processing of raw data. Concerning conventional sensors (RGs, THs, and FG), the data handling merely consists in (1) replacement of missing data or outliers, (2) resampling, and (3) conversion of hydrometric measurements into discharge values. Data reduction instead allows to retrieve rainfall intensity from CML-based raw data. Then, the semi-distributed hydrological model and its calibration/validation procedure are exhibited. Finally, we illustrate methods exploited for spatial interpolation of RG- and CML-based rainfall data.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

3.1 Data processing

Conventional sensors (rain gauges, thermometers, and flow gauge)

Raw data from RGs, THs and the FG were firstly processed to correct invalid measurements (missing data and outliers), which account for less than 1% in the period January 2018–June 2020. The process is different depending on the type of measurement. Invalid RG data were replaced interpolating valid observations from the nearby sensors by the IDW algorithm. Invalid TH data, as well as invalid FG measurements, were instead replaced by a linear interpolation. After data correction, the 10 min raw data were resampled to hourly time scale. Lastly, water level observations were converted into river discharge measurements by using the rating curves, validated by the Hydrographic Office of ARPA Lombardia.

Commercial Microwave Links

The CML raw data, collected by a network monitoring tool are minimum and maximum values of the transmitted and received power levels (TSL and RSL, respectively) every 15 min.

Microwave links of mobile networks are usually two-ways and provide dual-frequency operation, hence adding a certain degree of redundancy when it comes to rainfall estimates. Procedures for the conversion of RSL into rainfall rate have been detailed by several authors (e.g., Schleiss and Berne, 2010; Fenicia et al., 2012; Overeem et al., 2016) and are not discussed here extensively. The analysis of CML data in this work went through the following steps (Nebuloni et al., 2020b): (1) identification and removal of outliers and artifacts (i.e., occasional spikes which are not caused by rain); (2) classification of each 15 min time slot into dry or wet (i.e., rainy) by thresholding the difference between maximum and minimum RSL values; (3) estimation of the baseline, i.e., the RSL in absence of rain; (4) calculation of total signal attenuation as the difference between the baseline and the actual RSL; (5) identification and subtraction of the components of total attenuation not due to rainfall (e.g. wet antenna attenuation); (6) conversion of rain attenuation into average rain rate.

A few aspects of the above procedure deserve a discussion. First, the available RSL (and TSL) has a coarse 1 dB quantization step. That is, the time series of power (in W) have a random zero-mean error superimposed with rectangular distribution and limiting values equal to $\pm$ 12% of the measurement. It descends that it is impossible to distinguish between rain and quantization-induced noise below a certain rainfall intensity threshold. Figure 3 shows the minimum detectable rainfall intensity without ambiguity as a function of the CML path length with the CML frequency as parameter. The square markers correspond to the 50 CMLs in the study area divided in three groups according to their frequency. Continuous lines are drawn at four reference frequencies as well. Moreover, quantization affects the accuracy of rainfall intensity estimates. The accuracy of instantaneous measurements (at the 95% confidence level) is within 20% if the rainfall intensity exceeds 3 $\mathrm{mm\,h^{-1}}$ for the link with the most favorable combination between length and frequency. However, in the worst case, the above accuracy is achieved only if the rainfall intensity is above 10 $\mathrm{mm\,h^{-1}}$. The only way to mitigate quantization effects is to average in time.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

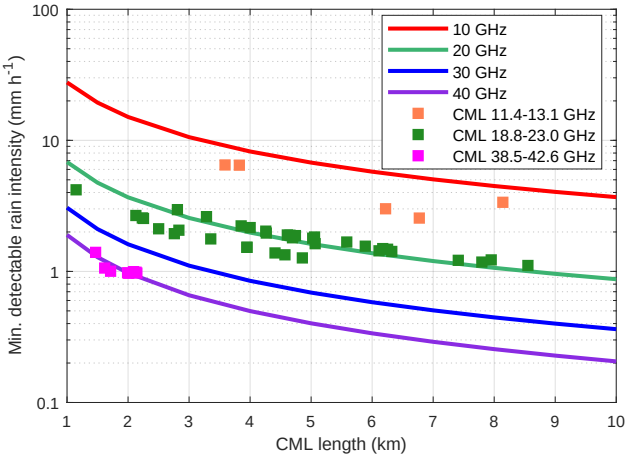

Figure 3. Minimum detectable rainfall intensity as a function of path length for link frequencies of 10, 20, 30 and 40 GHz, assuming a 1 dB quantization step on RSL. Squares represent the frequency-length of the 50 available CMLs in the study area.

Second, the relationship between rain attenuation per unit path length γ_R ($\mathrm{dB\,km^{-1}}$) and rainfall intensity R ($\mathrm{mm\,h^{-1}}$) is usually modelled by the following power-law function:

$$\gamma_R = \kappa R^{\alpha}, \tag{1}$$

where the coefficients κ and α have been tabulated by the International Telecommunication Union as a function of signal frequency and polarization (ITU-R P.838-3, 2005). In principle, the $\gamma_R - R$ relationship is dependent on the microphysics of rain as well, hence κ and α should be calibrated provided the characteristics of precipitation known in the climatic area where CMLs are deployed. In this work, raindrop size distribution data gathered from disdrometers were used to calculate the optimum value of κ and α coefficients following the procedure outlined in Luini et al. (2020).

Third, in the available CML data format, only the two extreme values of TSL and RSL are saved in every 15 min window. Therefore, if the average rainfall rate is to be estimated, for instance to calculate hourly accumulations, it is necessary to derive it from the extremes. To this aim, TSL and RSL time series sampled each 10 s were made available for a set of sample CMLs during the events considered here and processed as shown in Nebuloni et al. (2020a).

Fourth, it is assumed that rain attenuation measured over a CML of length L, is L times the attenuation per unit path length in Eq. (1), that is rain is considered uniform along the path. The effect of the inhomogeneity of precipitation can be relevant as CML paths range from about 1 km to nearly 9 km (Fig. 3). Some authors proposed to retrieve the spatial distribution of the rainfall field across the measurement area by processing all the CML data together, for instance through tomographic techniques. In this work, a simpler approach is used. Each CML is considered independently of the other and the corresponding rainfall measurement is given a weighting coefficient dependent on CML length, as discussed in Sec. 3.3.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

3.2 Hydrological model

We used a semi-distributed rainfall-runoff model, at hourly time scale. The catchment area is divided into 15 HRUs, which are considered meteorologically, geologically, and hydrologically homogeneous. The model parameters are hence set at HRU scale.

The river discharge at time t, $Q(t)$, in HRU's outlets is calculated as the sum of two main components

$$Q(t) = Q_s(t) + Q_g(t), \tag{2}$$

where Q_s is the contribution given by the surface runoff R^*, i.e., the portion of rainfall rejected by the soil, an Q_g is the groundwater discharge.

The computation of R^* (in mm) relies on the SCS-CN method (US Department of Agriculture Soil Conservation Service, 1985):

$$R^* = \frac{(P - I_a)^2}{P - I_a + S}, \tag{3}$$

where P (in mm) is the rain depth, S (in mm) is the maximum soil potential retention, and I_a (in mm) is the initial abstraction (calculated as a percentage, 20%, of S). According to USDA-SCS guidelines, the soil moisture condition antecedent to a storm event is classified depending on the value of the five-day antecedent rainfall. Here, we account for the actual soil moisture in a dynamical way, as proposed in the AnnAGNPS model (Bingner and Theurer, 2005) and also implemented in Ravazzani et al. (2007). In particular, the value of $S(t)$ is updated as a continuous function of the degree of soil saturation $\epsilon(t)$ (-)

$$S(t) = S_I \cdot \left\{ 1 - \left[\frac{\epsilon(t)}{\epsilon(t) + \exp(W_I - W_{II} \cdot \epsilon(t))} \right] \right\} \tag{4}$$

where the weights W_I (-) and W_{II} (-) are defined as follows

$$W_I = \ln\left[\frac{1}{1 - S_{III}/S_I} - 1 \right] + W_{II} \tag{5}$$

$$W_{II} = 2 \cdot \left[\ln\left(\frac{0.05}{1 - S_{II}/S_I} - 0.5 \right) - \ln\left(\frac{1}{1 - S_{III}/S_I} - 1 \right) \right] \tag{6}$$

being S_I, S_{II}, S_{III} the retention parameters associated with the curve numbers CN_I, CN_{II}, and CN_{III}, respectively. Finally, $\epsilon(t)$ is calculated as

$$\epsilon(t) = \frac{\theta(t) - \theta_{res}}{\theta_{sat} - \theta_{res}}, \tag{7}$$

where θ_{sat} (-) is the soil moisture at saturation conditions and θ_{res} (-) is the residual soil moisture.

To calculate Q_s at time t, the runoff is routed to the HRU's outlet, representing each HRU as a linear reservoir model (Dooge, 1973):

$$Q_s(t) = \int_0^t a \cdot r^*(\tau) \cdot A \cdot T_{lag}^{-1} \cdot \exp\left(-\frac{t - \tau}{T_{lag}} \right) d\tau, \tag{8}$$

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

where $a = 10^{-3}$ m mm^{-1} is a conversion factor, r^* is the surface runoff rate (in mm s^{-1}), A (in m^2) is the HRU area, and T_{lag} (in s) is the lag time calculated as 0.6 times the concentration time, T_c, for average natural watershed conditions and an approximately uniform distribution of runoff according to Mockus (1957) and de Simas (1996). The calculation of T_c, in each HRU, relies on the formula proposed by Ferro (2006).

The portion I (in mm) of total rainfall that infiltrates in the shallower layer of soil can either be lost by evapotranspiration, ET (in mm), or by percolation, D (in mm). Potential evapotranspiration (PET) is here calculated by the Hargreaves and Samani (1985) equation, which requires temperature data. The actual evapotranspiration (ET) is computed as a fraction of PET following Ravazzani et al. (2015). The water balance equation, referred to the shallower layer of soil with depth z (in mm) at time t, is formulated as

$$\theta(t) = \theta(t-1) + \frac{I(t-1) - D(t-1) - ET(t-1)}{z}, \tag{9}$$

where θ (-) is the actual soil moisture. $D(t)$ is the drainage flux calculated as

$$D(t) = c \cdot K_{sat} \cdot \epsilon(t)^{\frac{2+3B}{B}}, \tag{10}$$

where $c = 3.6 \cdot 10^6$ (in mm s m^{-1}) is a conversion factor, K_{sat} (in m s^{-1}) is the hydraulic conductivity at saturation, and B is the Brooks-Corey index (Brooks and Corey, 1964). Finally, $Q_g(t) = a \cdot \Delta T^{-1} \cdot D(t) \cdot A$, with $\Delta T = 3600$ s. The interaction among HRUs is represented by in series or in parallel reservoirs, according to the development of the river network, as exhibited in Fig. 1.b. CN values were taken from https://www.isprambiente.gov.it while θ_{res}, θ_{sat}, and B parameters were taken from Maidment (1993). K_{sat} and z are instead calibrated, as reported below.

Calibration and validation of the hydrological model

The hydrological model was calibrated using hourly rainfall depths from the RG network in Fig. 1.a. The chosen period for calibration is 1 January 2019–31 December 2019. We carried out a sensitivity analysis on model's parameters to identify the most important ones, to be adjusted with calibration, and we selected two of them, namely K_{sat} and z. The initial value of K_{sat}, for each HRU, was taken from Maidment (1993). Then, K_{sat} parameter was calibrated by applying a multiplicative coefficient. We tested, as multiplicative coefficients, several powers of 10, $\{10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}\}$. The range of variation for z was instead [10 cm, 3 m]. The parameters validation was carried out for a total of one year time interval, subdivided into two periods: the first is 1 July 2018–31 December 2018 and the second is 1 January 2020–30 June 2020. As model performance metric we selected the Nash and Sutcliffe (1970) efficiency, NSE. We identified the optimum combination of parameters through a trial and error calibration. The chosen parameters provided a NSE of 0.69 for the calibration and of 0.56 for the validation. We recall that the discharge simulation performances are generally considered reliable when NSE > 0.5 (Moriasi et al., 2007). It is also worth to notice that the calibration and validation steps were particularly troublesome due to the presence of the Cavo Diotti dam, which artificially regulates the outflow of Pusiano lake during flood events.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

3.3 Spatial interpolation of rainfall data

In the past, various methodologies have been tested and applied for spatial interpolation of CML retrieved rainfall measurements, considering different spatial scales (e.g., Fencl et al., 2013; Overeem et al., 2013; D'Amico et al., 2016; Haese et al., 2017; Chwala and Kunstmann, 2019; Graf et al., 2020; Eshel et al., 2021). Here, we exploited the simple and robust IDW method (Shepard, 1968) for both RG and CML measurements. However, since CMLs measurements are path-averaged we considered CMLs as point sensors collapsed in their midway point but we used an extra weight only to dump the effect of rainfall inhomogeneity associated to long links.

Following the IDW method, given n measurements $\{u(\mathbf{x}_1), ..., u(\mathbf{x}_n)\}$ at given points $\mathbf{x}_i$, with $i = 1, ..., n$ the interpolated value u, in $\mathbf{x}$, is calculated as

$$u(\mathbf{x}) = \begin{cases} \frac{\sum_{i=1}^{n} w_i' \cdot u(\mathbf{x}_i)}{\sum_{i=1}^{n} w_i'} & \text{if } d(\mathbf{x}_i, \mathbf{x}) \neq 0 \text{ for all } i; \\ u(\mathbf{x}_i) & \text{if } d(\mathbf{x}_i, \mathbf{x}) = 0 \text{ for some } i, \end{cases} \tag{11}$$

with

$$w_i' = \frac{1}{d(\mathbf{x}_i, \mathbf{x})^{\gamma}}, \tag{12}$$

where $d(\mathbf{x}_i, \mathbf{x})$ is the distance between the measuring point $\mathbf{x}_i$ and the coordinates of HRU's centroid $\mathbf{x}$, and $\gamma > 0$. The number of contributing measurements n is the one within a distance d_{max} from the query point. We identified appropriate values for the exponent γ and d_{max} by an iterative procedure presented herein. We estimated the precipitation, at each RG point, from the remainder of RGs at distance smaller than d_{max}, and calculated the Root Mean Square Error (RMSE) between observations and estimates. The process was repeated for several values of γ and d_{max}. We set a minimum of 10 km for d_{max}, to have at least one neighbour available for every considered RG. To this aim we exploited a larger set of 38 RGs (including the 13 RGs in Fig. 1.a) located on a wider area compared to only the Lambro basin and we used data from January 2018 to June 2020. The resulting RMSE distribution as a function of d_{max} and γ is reported in Fig. 4. We observe that the choice of d_{max} has a rather limited effect on the estimates, for $\gamma \geq 2$. The minimum RMSE is achieved when γ is slightly above 3. We therefore chose $\gamma = 3$. Finally, we selected $d_{max} = 10$ km, which provides the best RMSE when $\gamma = 3$.

To spatially interpolate CMLs rain rates, we took into account the CML length. Firstly, we handled CMLs as *virtual* RGs, considering the rain estimate as collapsed in the middle point of the CML length. Again we used IDW, this time with a weight given by w' multiplied by a length-dependent term w'' being

$$w'' = e^{-\beta^{-1} L}, \tag{13}$$

where L is the CML's length and β^{-1} the decreasing rate. In fact, the longer is the CML, the higher the probability that rainfall is unevenly distributed along its length. We chose $\beta = 5$, hence for $L = 8.6$ km (maximum CML path length), ω'' is equal to 0.18, while for $L = 1.1$ km (minimum path length), $\omega'' = 0.80$. Finally, as done in the case of RGs, to spatially interpolate CMLs data into each HRU centroid, we only considered CMLs whose middle points fall within $d_{max} = 10$ km from the HRU centroid.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

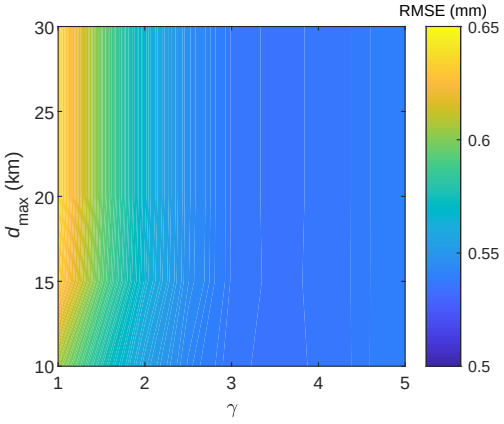

Figure 4. IDW calibration based on RGs: RMSE between observed and simulated rainfall depths, for different values of γ and d_{max}.

In addition of considering only RGs, or only CMLs, we accounted for the integration of RGs and CMLs measurements. In the following, we will refer to this option as CML+RG.

4 Results

The results of this paper are presented in the following two subsections. Sec. 4.1 provides a comparison on rainfall depths interpolated in the HRU centroids, by using data both from conventional and opportunistic sensors. Both accumulated rainfall values and hourly rainfall depths are considered at basin and sub-basin scale, for the 12 events. We therefore investigate whether there are some critical issues that might help to explain differences in the rainfall-runoff model outputs. Sec. 4.2 analyzes discharge performances, by comparing RG-, CML-, and RG+CML-driven simulations with the observed flow rates.

4.1 Comparison between RG and CML rainfall data in each HRU

Figure 5 shows the scatter plot of the areal-averaged rainfall depths accumulated at the end of each storm event. Yellow markers are CML against RG rainfall depths, while in orange are CML+RG against RG rainfall depths. The corresponding regression lines indicate a general good agreement between the two sets of sensors. For all the *Low rain rate* events (squares), estimates from CMLs (and CMLs+RGs) are lower than those from RGs. On the other side, CMLs tend to return higher estimates than RGs during *High rain rate* events (circles), even though the trend is not as evident.

We further investigated the above behaviour on the hourly time scale and on the sub-basin spatial scale by calculating the relative error, ΔE, of CML estimates, with respect to RG ones, for the hourly rain depths inferred in the 15 HRU centroids. We considered only wet hours (rain depth ≥ 1 mm for RG estimates in HRU centroids), that are 2061. Figure 6 shows a binned scatter plot (orange circles) of ΔE against RG-based rain depth, taken as the benchmark. The small circles in cyan are the

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

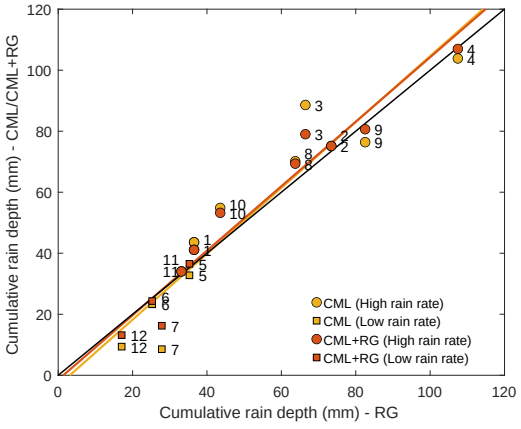

Figure 5. Rain depths, averaged over the catchment area and accumulated at the end of each event. The two different markers, circles and squares, respectively stand for *High rain rate* and *Low rain rate* events. The black line represents the 1:1 line of perfect matching between rain depth estimates from RGs and from CMLs (yellow) or from the combination of RGs and CMLs (orange). Yellow and orange lines are the corresponding regression lines.

2061 measured values. The x-axis has been divided into 11 bins. The x coordinate of each orange circle is the mid point of each bin interval, whereas the y coordinate is the mean of the ΔE values calculated over all the cyan circles falling within the bin. Finally, vertical bars represent the ± 1 standard deviation bounds. For rain depths smaller than about 3 mm, mean ΔE values are negative (CMLs estimates lower than RG estimates), due to the large numbers of zeroes in CML rain depths during wet hours (i.e. $\Delta E = -1$). Conversely, for rain rates higher than 3 mm, mean ΔE values are slightly higher than 0. Indeed, if the rain depth is smaller than 3 mm, only 30% of ΔE values falls in the range [-0.4, 0.4], whereas if it is larger than 3 mm, the percentage increases up to nearly 70%.

We therefore focused on the CML hourly wet-dry (see Sec. 2) classification, inferred in HRU centroids, again considering RG estimates as benchmark. A *false positive* occurs when an hourly slot is classified as wet by CMLs while it is dry according to RGs, and viceversa for a *false negative*. Figure 7 depicts a box plot of *false positive* and *false negative* percent occurrences. Events have been grouped according to their intensity. *Low rain rate* events clearly exhibit a higher median and a larger dispersion of *false negative* percentage than *High rain rate* events, whereas the occurrence of a *false positive* is relatively rare in both cases. These results confirm the inability of CMLs in detecting low rain rates, which depends on the quantization error issue discussed in Sec. 3.

Finally, in Fig. 8 we report box plots of ΔE values calculated for rain depths accumulated at the end of the 12 events, for each HRU. It does not emerge a general positive or negative trend and the median ΔE values are inside the range [-0.2,0,2] for 12 out of 15 HRUs. Moreover, in all the cases medians are lower than 0.3. Negative values of ΔE are mostly observed in HRUs 1, 2, 5, and 11. Such underestimations take place since CMLs, which have a non-negligible weight in the rainfall spatial

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

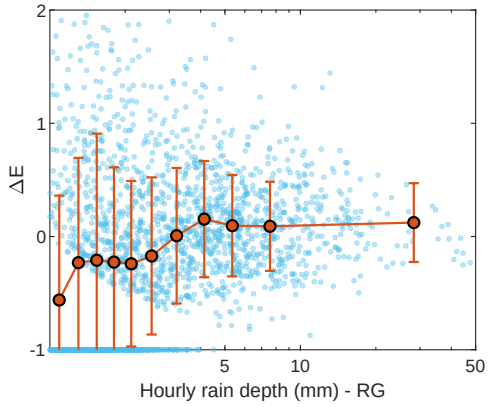

Figure 6. Relative difference ΔE between CML and RG hourly rain depths against RG rain depths. X-axis has a logarithmic scale.

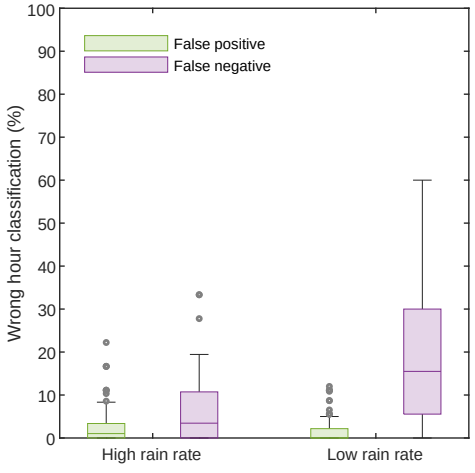

Figure 7. Percentage variation of hours subjected to wrong wet-dry classification (*false negative* or *false positive*), with respect to *High rain rate* and *Low rain rate* events.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

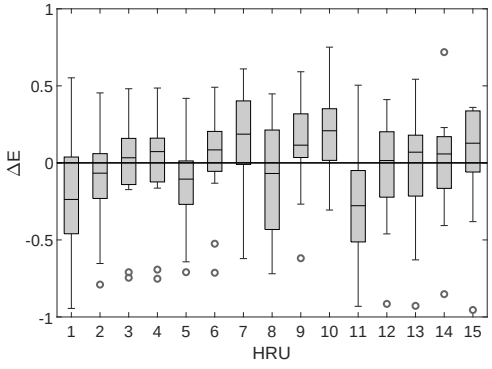

Figure 8. Box plots of ΔE for the 12 storm events grouped by HRU.

interpolation within these HRUs, are the least sensitive to low rain rates. In particular, the highest differences between RG and CML estimates are associated to HRU 11 where the two major contributing links are characterized by quite low frequencies ($\sim$ 18–19 GHz) and have lengths up to 6.3 km. We observed that such two CMLs systematically provide lower estimates, if compared to the closest RGs (analysis not reported here), with greater evidence for low rain rates hours.

4.2 Comparison between RG- and CML-driven discharge simulations

In this section, we investigate the matching between RG and CML focusing on discharge simulations at Lesmo river section. Again, the comparison is carried out for the 12 flood events selected. The model was fed by interpolated rainfall data from RGs, CMLs, and RGs+CMLs. The output performances were evaluated with three indices: (1) the Nash–Sutcliffe efficiency, NSE, (2) the relative error on peak discharge, REP, and (3) the relative error on flow volume, Dv. The formulations of these last two indices are respectively

$$\mathrm{REP} = \frac{Q_{sim}^{max} - Q_{obs}^{max}}{Q_{obs}^{max}}; \tag{14}$$

$$\mathrm{Dv} = \frac{V_{sim} - V_{obs}}{V_{obs}}, \tag{15}$$

where Q_{sim}^{max} is the simulated peak discharge, Q_{obs}^{max} is the observed peak discharge, V_{sim} is the simulated total flow volume, and V_{obs} is the observed total flow volume. Performances of the 12 discharge simulations, grouped by rainfall data input, are summarized in Fig. 9, through box plots. The statistical dispersion (represented by the interquartile range, IQR) of CML-based discharge simulations are larger than RG-based simulations. The use of CML interpolated data into the rainfall-runoff model seems to produce higher uncertainty, with respect to RG data. The combined use of RGs and CMLs instead decreases the statistical dispersion of results and leads to performances closer to those achieved through RG. Generally, CMLs exhibit worse

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

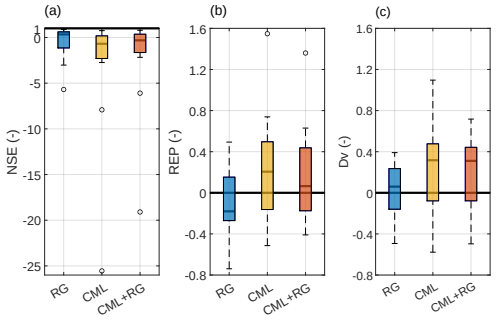

Figure 9. From panel (a) to (c), box plots of performance metrics, namely NSE, REP, and Dv, for the 12 selected flood events. The optimum values correspond to the bold black lines.

performance than RGs in terms of NSE and Dv. As for REP, the two set of sensors produce comparable errors of opposite sign, hence their combined use leads to an optimum value of the median error (0.06).

Figure 10 shows a scatter plot of observed against simulated flow peaks. Firstly, it shows that the best match between observations and simulations is not always achieved by RG-based simulations. In fact, for events 1, 3, 5, and 11 the optimum matching is given from CML- or CML+RG-based simulations. Moreover, the *Low rain rate* events typically result in underestimated peak flow simulations with respect to the observations, considering either conventional or unconventional sensors, with an exception for event 5. Figure 11 reports model inputs and outputs for event 5 and 2. The first event is an autumn stratiform event, characterized by low rain intensity. In Fig. 11.a we can see that CML-based estimates in HRUs 1, 2, 5, 8, and 11 are quite low, with respect to RGs, due to the difficulties of relevant CMLs in detecting light rain, as also highlighted in Fig. 8. In contrast, in the southern HRUs 10, 12, 13, 14, and 15, which have much more influence in the generation of river discharge, CML estimates are higher than RG ones. Finally, Fig. 11.b shows that the CML-driven simulation better represents the observed outflow hydrograph, with respect to the RG-driven simulation. In particular, the best performance is gained when both the two types of rainfall sensors are used, and it provides an excellent Dv, equal to 0.03. The most problematic CMLs are mostly located in the northern portion of the basin and have less impact in generating discharge. Event 2 is instead a typical intense convective summer event, characterized by a single rainfall peak. As rain rates are high all over the basin, contrary to event 5, we do not observe lower CML estimates with respect to RG estimates, for HRU 1, 2, and 5. However, lower values are still present in HRU 11. In general, Fig. 11.c reveals a good agreement between RG and CML estimates. River discharge simulations, reported in Fig. 11.d, are satisfactory, considering all the 3 input data. NSE values obtained respectively from RGs, CMLs, and RGs+CMLs data are 0.86, 0.77, and 0.80.

As the hydrological model has been calibrated with RG-detected rainfall data, it can be easily guessed that the best model performances are mostly achieved with RG data as input, as well. In fact, the major drawback of the present work is definitely that we did not rely on a large and real-time CML-based dataset and a calibration similar to that implemented with RG data cannot be carried out. Nevertheless, we tried to overcome this problem by recalibrating model parameters, with CML and

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

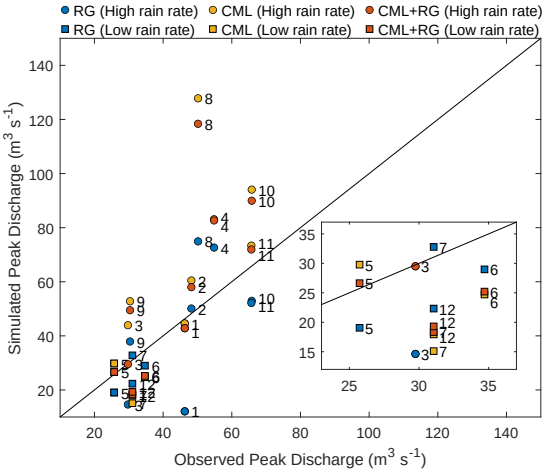

Figure 10. Observed peak discharge against simulated one from RG (blue), CML (yellow), and CML+RG (orange). The two different markers, circles and squares, respectively stand for *High rain rate* and *Low rain rate* events. Inset figure reports a zoom for events with a low peak discharge.

CML+RG rainfall estimates as input, relying on the 12 available flood events. We considered as optimum parameter values those providing the highest median NSE values. Performance indices, subdivided by type of rainfall data input and type of calibration, are summarized in Fig. 12 by box plots. Please note that CML- and CML+RG-based calibration improves the performance of the model when fed by unconventional input data. In particular, NSE values are comparable with those achieved by the use of RG data with a RG-based calibration. In fact, median NSE values for RG inputs with RG-based calibration, CML inputs with CML-based calibration, and CML+RG inputs with CML+RG-based calibration are 0.34, 0.35, and 0.38, respectively. For REP values we generally observe underestimations of observed peak flow, considering CML- and CML+RG calibration but smaller IQR when compared with RG-based calibration. Concerning Dv values, performances for the CML- and CML+RG-based calibration are quite satisfactory, despite the combination providing the best performance is still RG inputs with RG-based calibration.

5 Discussion

The analysis on interpolated rainfall data carried out in Sec. 4.1 reveals that CML and RG estimates of accumulated areal-averaged rainfall depths are comparable. However, two issues emerged.

First, CMLs exhibit a different behaviour depending on event's intensity, since their sensitivity varies with length and frequency. Specifically, they return lower values of rainfall depth and rainfall accumulation in correspondence of *Low rain rate* events, both at the basin and sub-basin scale. In fact, due to a coarse quantization of the raw data, CMLs are not sensitive to low rain intensity, hence, when the rainfall depth value over a certain lapse of time is required (in our case 1 h), this limitation may

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

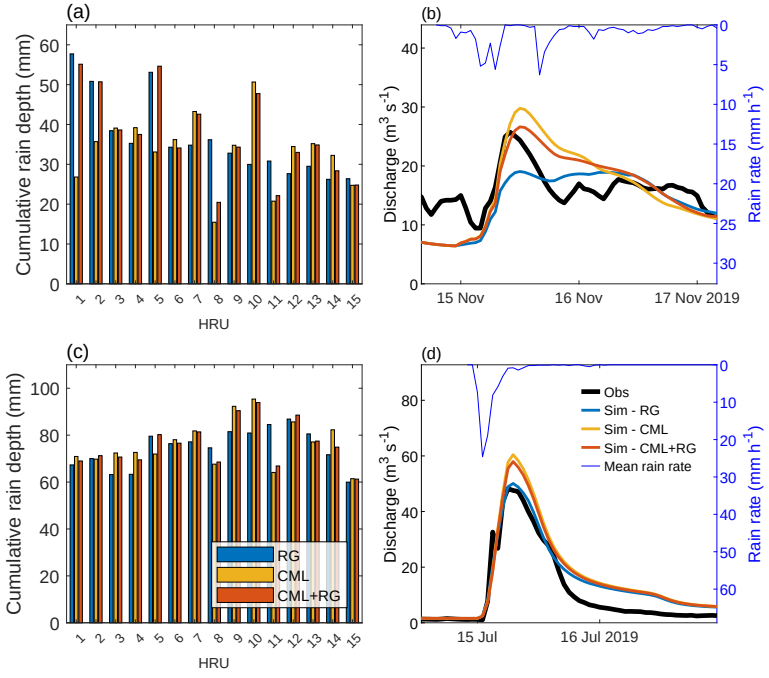

Figure 11. In panel (a) and (c), cumulative rain depth estimates from RG, CML, and CML+RG for storm/flood events 5 and 2, respectively. In panel (b) and (d) discharge observations and simulations gathered by using the three different input data, for the two same events.

lead to large errors, especially when light rain goes on for a long time. We instead observed a good agreement between CML- and RG-based estimates for high rain rates. The problem of CML underestimation mainly involves northern HRUs (specifically HRUs 1, 2, 5, and 11), since are covered by less sensitive CMLs. However, we observed that these underestimations do not strongly impact the river discharge simulations since the northernmost HRUs influence less the runoff generation in the outlet section.

The second issue is the different CML density over the HRUs. It is well known that spatial interpolation methods are sensitive to sensors' density (Xu et al., 2013), and consequently the relatively large distance of the available CMLs from the HRU centroids in the most scarcely populated areas (northern HRUs) may lead to loss of reliability of estimated rainfall depths. However, such an aspect appears to be less relevant when compared with the first one. In fact, a comparison of Fig. 2 and Fig. 8 does not show an evident correlation between the mean CMLs-centroids distance and ΔE.

Last but not least, model performances are influenced by the calibration process of model parameters. Similar model performances, in terms of NSE index, can be achieved with all the three types of input data (RGs, CMLs, and CMLs+RGs) if the calibration is carried out with the respective data inputs. This means that, after a proper calibration, opportunistic sensors could

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

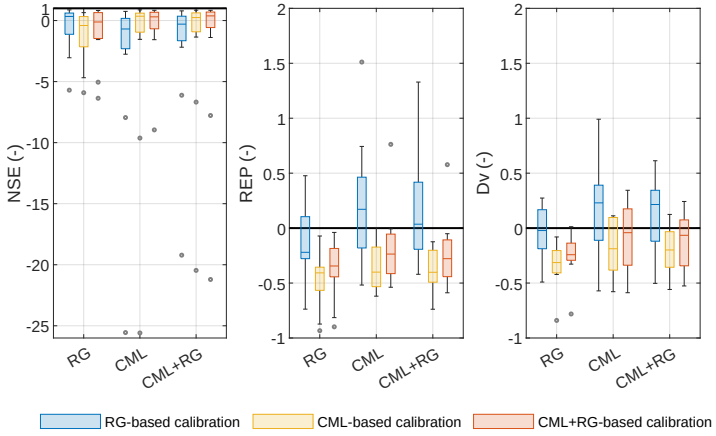

Figure 12. Box plots of the performance indices for the 12 flood events obtained with three different calibration sets (drawn in as many colours) and with three different input types (on the x-axis).

be exploited in semi-distributed hydrological models, as well as RGs. In particular we found that, after calibration, the set RGs+CMLs is the one providing the highest median NSE. However, it is worth to highlight that we calibrated the parameters on the basis of only 12 flood events. In order to assess a robust calibration and the associated validation, a larger dataset of CML-based rainfall events should be processed.

6 Conclusions

In this work, we assessed the use of CMLs as opportunistic rainfall sensors within the hydrological modelling. We focused on Lambro, a peri-urban catchment, 260 km^2 in area, located north of Milan (Italy) and covered by 50 CMLs that are part of the network owned by a major mobile operator. Lambro's area is covered by 13 RGs as well, which we used both as an independent rainfall dataset and in combination with CMLs. We implemented a semi-distributed hydrological model and carried out two types of comparison between CML and RG data. First, we considered rainfall data (hourly rainfall depths, the input of the hydrological model, and total accumulations at the storm end, for a sample of 12 storm events) interpolated at the HRU centroids. Then, we compared river discharge simulations (model output) from RGs, CMLs and RGs+CMLs against flow measurements.

Concerning the comparison on rainfall data, we found out that high intensity events detected by CMLs are in accordance with RG measurements. On the other hand we came across a critical aspect, which is the inability of CMLs to detect low rain rates, due to the coarse 1 dB quantization step of raw data (i.e. received power levels). The minimum detectable rainfall intensity depends on the operation frequency of CMLs as well as on their length, and, for the available set of CMLs, it ranges from 3

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021


Hydrology and Earth System Sciences
Discussions

$\mathrm{mm\,h^{-1}}$ to 10 $\mathrm{mm\,h^{-1}}$. Such a limitation results in the underestimation of rain depths interpolated in the HRU centroids for low intensity storm events, when compared from RG based rainfall data.

The hydrographs simulated by the hydrological model highlight better performances in terms of NSE and Dv metrics in the case of RGs rather than of CMLs. This result is not surprising as the model was calibrated using RG data throughout one year data. Nevertheless, satisfactory REP values are achieved through the use of CML and CML+RG data as inputs into the RG-based calibrated model. We notably observed that the underestimating behaviour of CMLs during low intensity events does not impact much discharge simulations, in fact peak discharges simulated from CMLs are not systematically lower than RG-driven simulations. This happening since the least sensitive CMLs are located over the norther HRUs.

By calibrating the model with CML data and by using the same as input, it is possible to improve the model performance which becomes comparable with the case of RG-calibration and RG-input. Even a slightly better performance can be gathered with a CML+RG-based calibration and CML+RG data as input.

An extension of the CML-based dataset of events would bring great benefits to the present work. Firstly, it would allow the development of a more robust statistical analysis on storm/flood events. Secondly, it would enable a proper calibration and validation of the hydrological model based on CML data as rainfall input.

To enhance this work, it would also be useful to implement a CML-driven distributed model, which is expected to provide a more accurate description of the spatial variability of the precipitation field with respect to a semi-distributed one. In such a case, the CML measurements would be better exploited by the use of advanced methods for spatial reconstruction of the rainfall field. For instance, techniques as the tomographic reconstruction algorithm (D'Amico et al., 2016) or the stochastic reconstruction based on copulas (Haese et al., 2017; Salvadori et al., 2007), take advantage of the path integrated nature of CML measurements.

Finally, it is worth to notice that, although we showed that CML rainfall data can be successfully assimilated into hydrological models, their integration into real-time operational platforms (e.g. early warning systems) remains challenging. A number of aspects should be considered including:

- generation of CML raw data formats suitable for rainfall estimation;
- real-time collection of raw data, which should be transparent to network operation;
- data transfer to a control centre;
- heavy data reduction process, especially if large sets of CMLs are managed.

The above mentioned issues suggest a systematic cooperation with mobile operators, who are the owners of CML network infrastructure.

Author contributions. C.D.M., R.N., M.D.A., A.G. made the conceptualization; R.N. handled CML data; G.C. handled gauge data, G.C., C.D., C.D.M. developed the hydrological model; G.C. made the model calibrations and validations; G.C. prepared Figs. 1-2, 4-12; R.N. prepared Fig.3; G.C. wrote the first draft of the manuscript; all the authors reviewed the manuscript.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

Competing interests. All authors declare that they have no conflicts of interest.

Acknowledgements. We wish to acknowledge the support by Fondazione CARIPLO through the MOPRAM project (http://www.mopram.it).
We thank ARPA Lombardia for providing us with meteorological and hydrological data.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

References

Alberoni, P., Andersson, T., Mezzasalma, P., Michelson, D., and Nanni, S.: Use of the vertical reflectivity profile for identification of anomalous propagation, Meteorol. Appl., 8, 257–266, https://doi.org/10.1017/S1350482701003012, 2001.

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., and Javelle, P.: Sensitivity of hydrological models to uncertainty in rainfall input, Hydrol. Sci. J., 56, 397–410, https://doi.org/10.1080/02626667.2011.563742, 2011.

Bárdossy, A. and Das, T.: Influence of rainfall observation network on model calibration and application, Hydrol. Earth Syst. Sci., 12, 77–89, https://doi.org/10.5194/hess-12-77-2008, 2008.

Bengtsson, L.: Daily and hourly rainfall distribution in space and time–conditions in southern Sweden, Hydrol. Res., 42, 86–94, https://doi.org/10.2166/nh.2011.080b, 2011.

Bingner, R. L. and Theurer, F. D.: AnnAGNPS: technical description, Tech. rep., US Dept Agriculture, Agricultural Research Services, 2005.

Brauer, C. C., Overeem, A., Leijnse, H., and Uijlenhoet, R.: The effect of differences between rainfall measurement techniques on groundwater and discharge simulations in a lowland catchment, Hydrol. Processes, 30, 3885–3900, https://doi.org/10.1002/hyp.10898, 2016.

Brooks, R. H. and Corey, A. T.: Hydraulic properties of porous media, Fort Collins, Colorado State University, CO, USA, 1964.

Chen, D., Ou, T., Gong, L., Xu, C.-Y., Li, W., Ho, C.-H., and Qian, W.: Spatial interpolation of daily precipitation in China: 1951–2005, Adv. Atmos. Sci., 27, 1221–1232, https://doi.org/10.1007/s00376-010-9151-y, 2010.

Chwala, C. and Kunstmann, H.: Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges, Wiley Interdiscip. Rev.: Water, 6, e1337, https://doi.org/10.1002/wat2.1337, 2019.

Cugerone, K. and De Michele, C.: Johnson SB as general functional form for raindrop size distribution, Water Resour. Res., 51, 6276–6289, https://doi.org/10.1002/2014WR016484, 2015.

Dawdy, D. R. and Bergmann, J. M.: Effect of rainfall variability on streamflow simulation, Water Resour. Res., 5, 958–966, https://doi.org/10.1029/WR005i005p00958, 1969.

de Simas, M. J. C.: Lag-time characteristics in small watersheds in the United States, UMI, Ann Arbor, MI, USA, 1996.

Delhomme, J. P.: Kriging in the hydrosciences, Adv. Water Resour., 1, 251–266, https://doi.org/10.1016/0309-1708(78)90039-8, 1978.

Dooge, J.: Linear theory of hydrologic systems, 1468, Agricultural Research Service, US Department of Agriculture, 1973.

D'Amico, M., Manzoni, A., and Solazzi, G. L.: Use of operational microwave link measurements for the tomographic reconstruction of 2-D maps of accumulated rainfall, IEEE Geosci. Remote Sens. Lett., 13, 1827–1831, https://doi.org/10.1109/LGRS.2016.2614326, 2016.

Eshel, A., Messer, H., Kunstmann, H., Alpert, P., and Chwala, C.: Quantitative analysis of the performance of spatial interpolation methods for rainfall estimation using commercial microwave links, J. Hydrometeorol., 22, 831–843, https://doi.org/10.1175/JHM-D-20-0164.1, 2021.

EU Water Directors: Best Practices on flood prevention, protection and mitigation, in: Meetings in Budapest on 30 November and 1 December 2002, and in Bonn on 5/6 February 2003., http://ec.europa.eu/environment/water/flood_risk/pdf/flooding_bestpractice.pdf, 2003.

Fencl, M., Rieckermann, J., Schleiss, M., Stránskỳ, D., and Bareš, V.: Assessing the potential of using telecommunication microwave links in urban drainage modelling, Water Sci. Technol., 68, 1810–1818, https://doi.org/IWA 10.2166/wst.2013.429, 2013.

Fenicia, F., Pfister, L., Kavetski, D., Matgen, P., Iffly, J.-F., Hoffmann, L., and Uijlenhoet, R.: Microwave links for rainfall estimation in an urban environment: Insights from an experimental setup in Luxembourg-City, J. Hydrol., 464–465, 69–78, https://doi.org/10.1016/j.jhydrol.2012.06.047, 2012.

Ferro, V.: Riqualificazione ambientale dei corsi d'acqua, Quaderni di Idronomia Montana, Nuova Editoriale Bios, 2006.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Hydrology and Earth System Sciences
Discussions

Giuli, D., Toccafondi, A., Gentili, G. B., and Freni, A.: Tomographic reconstruction of rainfall fields through microwave attenuation measurements, J. Appl. Meteorol. Climatol., 30, 1323–1340, https://doi.org/10.1175/1520-0450(1991)030<1323:TRORFT>2.0.CO;2, 1991.

Graf, M., Chwala, C., Polz, J., and Kunstmann, H.: Rainfall estimation from a German-wide commercial microwave link network: optimized processing and validation for 1 year of data, Hydrol. Earth Syst. Sci., 24, 2931–2950, https://doi.org/10.5194/hess-24-2931-2020, 2020.

Haese, B., Hörning, S., Chwala, C., Bárdossy, A., Schalge, B., and Kunstmann, H.: Stochastic reconstruction and interpolation of precipitation fields using combined information of commercial microwave links and rain gauges, Water Resour. Res., 53, 10 740–10 756, https://doi.org/10.1002/2017WR021015, 2017.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Appl. Eng. Agr., 1, 96–99, https://doi.org/10.13031/2013.26773, 1985.

Ignaccolo, M. and De Michele, C.: One, No One, and One Hundred Thousand: The Paradigm of the Z–R Relationship, J. Hydrometeorol., 21, 1161–1169, https://doi.org/10.1175/JHM-D-19-0177.1, 2020.

ITU-R P.838-3: Specific attenuation model for rain for use in prediction methods, Tech. rep., 2005.

Jaffrain, J., Studzinski, A., and Berne, A.: A network of disdrometers to quantify the small-scale variability of the raindrop size distribution, Water Resour. Res., 47, https://doi.org/10.1029/2010WR009872, 2011.

Jameson, A. R. and Kostinski, A.: Spurious power–law relations among rainfall and radar parameters, Q. J. R. Meteorolog. Soc., 128, 2045–2058, https://doi.org/10.1256/003590002320603520, 2002.

Kidd, C. and Huffman, G.: Global precipitation measurement, Meteorol. Appl., 18, 334–353, https://doi.org/10.1002/met.284, 2011.

Köppen, W. P.: Die Klimate der Erde: Grundriss der Klimakunde, Geogr. J., 65, 1925.

Lombardi, G., Ceppi, A., Ravazzani, G., Davolio, S., and Mancini, M.: From Deterministic to Probabilistic Forecasts: The 'Shift-Target' Approach in the Milan Urban Area (Northern Italy), Geosciences, 8, 181, https://doi.org/10.3390/geosciences8050181, 2018.

Luini, L., Roveda, G., Zaffaroni, M., Costa, M., and Riva, C.: The Impact of Rain on Short *E*-Band Radio Links for 5G Mobile Systems: Experimental Results and Prediction Models, IEEE Trans. Antennas Propag., 68, 3124–3134, https://doi.org/10.1109/TAP.2019.2957116, 2020.

Maidment, D.: Handbook of hydrology, McGraw-Hill, New York, NY, USA, 1993.

Marshall, J. S. and Palmer, W. M.: The distribution of raindrops with size, J. Meteorol., 5, 165–166, https://doi.org/10.1175/1520-0469(1948)005, 1948.

Masseroni, D., Cislaghi, A., Camici, S., Massari, C., and Brocca, L.: A reliable rainfall–runoff model for flood forecasting: review and application to a semi-urbanized watershed at high flood risk in Italy, Hydrol. Res., 48, 726–740, https://doi.org/10.2166/nh.2016.037, 2017.

Messer, H., Zinevich, A., and Alpert, P.: Environmental monitoring by wireless communication networks, Science, 312, 713–713, https://doi.org/10.1126/science.1120034, 2006.

Mockus, V.: Use of storm and watershed characteristics in synthetic unit hydrograph analysis and application, US Soil Conservation Service, 1957.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Transactions of the ASABE, 50, 885–900, https://doi.org/10.13031/2013.23153, 2007.

Myers, D. E.: Co-kriging—New developments, in: Geostatistics for natural resources characterization, pp. 295–305, Springer, 1984.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nebuloni, R., D'Amico, M., Cazzaniga, G., and De Michele, C.: On the Use of Minimum and Maximum Attenuation for Retrieving Rainfall Intensity Through Commercial Microwave Links, URSI Radio Sci. Lett., 2, https://doi.org/10.46620/20-0062, 2020a.

Nebuloni, R., Tagliaferri, D., D'Amico, M., Sileo, A., Cazzaniga, G., and Deidda, C.: Rainfall Retrieval Through Commercial Microwave Links in Valmalenco (North Italy), in: 2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science, pp. 1–4, https://doi.org/10.23919/URSIGASS49373.2020.9232227, 2020b.

New, M., Todd, M., Hulme, M., and Jones, P.: Precipitation measurements and trends in the twentieth century, Int. J. Climatol., 21, 1889–1922, https://doi.org/10.1002/joc.680, 2001.

Obled, C., Wendling, J., and Beven, K.: The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data, J. Hydrol., 159, 305–333, https://doi.org/10.1016/0022-1694(94)90263-1, 1994.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, Proc. Natl. Acad. Sci., 110, 2741–2745, https://doi.org/10.1073/pnas.1217961110, 2013.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network, Atmospheric Measurement Techniques, 9, 2425–2444, https://doi.org/10.5194/amt-9-2425-2016, https://amt.copernicus.org/articles/9/2425/2016/, 2016.

Parkes, B., Wetterhall, F., Pappenberger, F., He, Y., Malamud, B., and Cloke, H.: Assessment of a 1-hour gridded precipitation dataset to drive a hydrological model: a case study of the summer 2007 floods in the Upper Severn, UK, Hydrol. Res., 44, 89–105, https://doi.org/10.2166/nh.2011.025, 2013.

Pastorek, J., Fencl, M., Rieckermann, J., and Bareš, V.: Commercial microwave links for urban drainage modelling: The effect of link characteristics and their position on runoff simulations, J. Environ. Manage., 251, 109 522, https://doi.org/10.1016/j.jenvman.2019.109522, 2019.

Raghavan, S.: Radar Meteorology, Springer, 2013.

Rahimi, A., Holt, A., Upton, G., and Cummings, R.: Use of dual-frequency microwave links for measuring path-averaged rainfall, Journal of Geophysical Research: Atmospheres, 108, 4467, https://doi.org/10.1029/2002JD003202, 2003.

Rauber, R. M. and Nesbitt, S. W.: Radar meteorology: A first course, John Wiley & Sons, Glasgow, UK, 2018.

Ravazzani, G., Mancini, M., Giudici, I., and Amadio, P.: Effects of soil moisture parameterization on a real-time flood forecasting system based on rainfall thresholds, in: Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management, Proc. Symposium HS 2004 at IUGG 2007, pp. 407—-416, IAHS Publ., Perugia, 2007.

Ravazzani, G., Bocchiola, D., Groppelli, B., Soncini, A., Rulli, M. C., Colombo, F., Mancini, M., and Rosso, R.: Continuous streamflow simulation for index flood estimation in an Alpine basin of northern Italy, Hydrological Sciences Journal, 60, 1013–1025, https://doi.org/10.1080/02626667.2014.916405, 2015.

Ravazzani, G., Amengual, A., Ceppi, A., Homar, V., Romero, R., Lombardi, G., and Mancini, M.: Potentialities of ensemble strategies for flood forecasting over the Milano urban area, J. Hydrol., 539, 237–253, https://doi.org/10.1016/j.jhydrol.2016.05.023, 2016.

Roversi, G., Alberoni, P. P., Fornasiero, A., and Porcù, F.: Commercial microwave links as a tool for operational rainfall monitoring in Northern Italy, Atmos. Meas. Tech., 13, 5779–5797, https://doi.org/10.5194/amt-13-5779-2020, 2020.

Salvadori, G., De Michele, C., Kottegoda, N. T., and Rosso, R.: Extremes in nature: an approach using copulas, vol. 56, Springer, Dordrecht, The Netherlands, 2007.

https://doi.org/10.5194/hess-2021-389
Preprint. Discussion started: 22 July 2021

Schleiss, M. and Berne, A.: Identification of Dry and Rainy Periods Using Telecommunication Microwave Links, IEEE Geosci. Remote Sens. Lett., 7, 611–615, https://doi.org/10.1109/LGRS.2010.2043052, 2010.

Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 1968 23rd ACM national conference, pp. 517–524, ACM, New York, NY, USA, https://doi.org/10.1145/800186.810616, 1968.

Smiatek, G., Keis, F., Chwala, C., Fersch, B., and Kunstmann, H.: Potential of commercial microwave link network derived rainfall for river runoff simulations, Environ. Res. Lett., 12, 034 026, https://doi.org/10.1088/1748-9326/aa5f46/meta, 2017.

Stransky, D., Fencl, M., and Bares, V.: Runoff prediction using rainfall data from microwave links: Tabor case study, Water Sci. Technol., 2017, 351–359, https://doi.org/10.2166/wst.2018.149, 2018.

Thiessen, A. H.: Precipitation averages for large areas, Monthly weather review, 39, 1082–1089, https://doi.org/10.1175/1520-0493(1911)39<1082b:PAFLA>2.0.CO;2, 1911.

US Department of Agriculture Soil Conservation Service: National Engineering Handbook, section 4 - Hydrology, Washington, DC, USA, 1985.

Xie, P., Rudolf, B., Schneider, U., and Arkin, P. A.: Gauge-based monthly analysis of global land precipitation from 1971 to 1994, J. Geophys. Res.: Atmos., 101, 19 023–19 034, https://doi.org/10.1029/96JD01553, 1996.

Xu, H., Xu, C.-Y., Chen, H., Zhang, Z., and Li, L.: Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China, J. Hydrol., 505, 1–12, https://doi.org/10.1016/j.jhydrol.2013.09.004, 2013.

Younger, P. M., Freer, J. E., and Beven, K. J.: Detecting the effects of spatial variability of rainfall on hydrological modelling within an uncertainty analysis framework, Hydrol. Processes, 23, 1988–2003, https://doi.org/10.1002/hyp.7341, 2009.