

Improved input to distributed hydrologic model in areas with sparse subdaily rainfall data using multivariate daily rainfall disaggregation

Marija Ivković, Andrijana Todorović and Jasna Plavšić

ABSTRACT

Flood forecasting relies on good quality of observed and forecasted rainfall. In Serbia, the recording rain gauge network is sparse and rainfall data mainly come from dense non-recording rain gauges. This is not beneficial for flood forecasting in smaller catchments and short-duration events, when hydrologic models operating on subdaily scale are applied. Moreover, differences in rainfall amounts from two types of gauges can be considerable, which is common in operational hydrological practice. This paper examines the possibility of including daily rainfall data from dense observation networks in flood forecasting based on subdaily data, using the extreme flood event in the Kolubara catchment in May 2014 as a case study. Daily rainfall from a dense observation network is disaggregated to hourly scale using the MuDRain multivariate disaggregation software.

The disaggregation procedure results in well-reproduced rainfall dynamics and adjusts rainfall volume to the values from the non-recording gauges. The fully distributed wflow_hbv model, which is under development as a forecasting tool for the Kolubara catchment, is used for flood simulations with two alternative hourly rainfall data. The results show an improvement when the disaggregated rainfall from denser network is used, thus indicating the significance of better representation of rainfall temporal and spatial variability for flood forecasting.

Key words | flood forecasting, HBV model, rainfall-runoff model, rainfall spatial distribution, sparse network, subdaily rainfall disaggregation

Marija Ivković (corresponding author)
Republic Hydrometeorological Service of Serbia,
Kneza Višeslava 66, Belgrade,
Serbia
E-mail: marija.ivkovic@hidmet.gov.rs

Andrijana Todorović
Jasna Plavšić
University of Belgrade – Faculty of Civil
Engineering,
Bulevar kralja Aleksandra 73, Belgrade,
Serbia

INTRODUCTION

Flood forecasting is a very important segment of water resources management. The accuracy of the forecasts depends on the meteorological input and a hydrologic model (Pappenberger *et al.* 2005). Accurate flood flow forecasts require well-estimated initial conditions in catchments (primarily initial soil moisture content) and good precipitation forecasts. Application of fully distributed models for flood forecasting is recommended because these models can accommodate spatial variability of the processes and consequently simulate different spatial origin of runoff, which determines the response of a basin (Pechlivanidis *et al.* 2017). The time when the flood wave starts to rise

and the time of peak occurrence in a river section are very important for flood forecasting. If precipitation or temperature spatial and temporal distributions are not well represented, the simulated flood hydrograph can be delayed or be ahead of the one observed.

Precipitation controls most hydrologic processes. Rainfall varies in both space and time leading to variations in catchment response (Singh 1997) and accurate assessment of rainfall spatial distribution is an important part of the catchment modelling. Heavy rainfall on a small part of the catchment and moderate rainfall over the whole catchment can both cause floods, but with different hydrographs. Storm

movement also affects the shape of flood hydrographs (Singh 1997). Due to the complexity of the problem, there is still no general conclusion about the relationship between rainfall spatial variability and catchment response (Arnaud *et al.* 2002) and the effects of rainfall variability are usually assessed indirectly using the hydrologic models (Obled *et al.* 1994; Segond *et al.* 2007).

Rainfall input is recognized as a source of considerable uncertainty in hydrological modelling (Sangati & Borga 2009; Kretschmar *et al.* 2016). Several studies (Obled *et al.* 1994; Arnaud *et al.* 2002) have shown that rainfall spatial variability significantly influences the ability of hydrologic models to predict runoff volume, peak flows and time to peak. Hydrological simulations with spatially distributed rainfall data are preferred to the basin-lumped data (e.g., Girons Lopez & Seibert 2016). Andreassian *et al.* (2001) demonstrated that high quality rainfall input yields better and more consistent model performance and Vaze *et al.* (2011) reported increased model efficiency with increasing rainfall spatial resolution. It is therefore of primary importance for hydrological modelling that the assessment of rainfall spatial distribution is as accurate as possible.

Hydrologic response also depends on rainfall temporal variation. Flood events in smaller catchments may occur within a few hours when the intensities are highly variable (Dolšak *et al.* 2016) and therefore the flood forecasting systems should rely on rainfall data of fine temporal resolution. In most parts of Serbia, subdaily rainfall data (as well as other meteorological data) are very limited, while daily data are available at a sufficient number of sites with non-recording rain gauges. The flood events in most of the flood-prone areas occur on a subdaily scale and the flood prediction systems cannot benefit from the hydrologic simulations with such a time resolution. In addition, the majority of the recording rain gauges do not measure precipitation during winter and the hydrologic model is not able to simulate snow accumulation or reproduce the volume of the spring snowmelt. There are also uncertainties in the observations from recording gauges due to, for example, high rainfall intensities (e.g., Maksimović *et al.* 1991).

Spatially distributed rainfall can be introduced to hydrologic modelling and forecasting in different ways. Radar-based rainfall spatial distribution is one of the possible solutions (e.g., Guardiola-Albert *et al.* 2016), but this approach

can be subject to great uncertainties if the radar data are not carefully applied (Einfalt *et al.* 2004; Thorndahl *et al.* 2017). Another way to extract information from spatially distributed daily rainfall data is to apply one of many disaggregation techniques appropriate for application in hydrology. The methods suitable for disaggregation of rainfall data range from simple to complex. The simplest technique is to create a uniform temporal distribution of daily rainfall, but the credibility of this approach may be questionable in case of localized intensive rainfall such as convective storms (Gutierrez-Magness & McCuen 2004). A simple proportional adjustment method commonly used consists of normalizing rainfall mass curves from a nearby recording rain gauge and transferring it to the site of interest by multiplying with its daily rainfall depth (Blöschl & Sivapalan 1995). More complex methods include multivariate disaggregation methods and stochastic generation of subdaily rainfall data (e.g., Abdellatif *et al.* 2013).

Univariate temporal disaggregation methods produce synthetic subdaily series consistent with the available daily time series and statistically consistent with the hourly time series, while rainfall depths and the maximum intensity are arbitrarily distributed within the day. The multivariate disaggregation method proposed by Koutsoyiannis *et al.* (2003) uses limited observations of sub-daily rainfall and extracts additional information from daily time series from the non-recording rain gauges to create spatially and temporally consistent subdaily series at locations of non-recording rain gauges. The model can also properly reproduce dry intervals, which is very important for rainfall disaggregation (e.g., Sivakumar 2005). This approach has practical significance if subdaily data exist within a region of interest and if there is a significant cross-correlation among the stations.

Development of the flood forecasting system for the Kolubara River in Serbia based on the fully distributed wflow_hbv model (Schellekens 2013) with hourly meteorological input started in April 2014, when the first automatic rain gauges were installed within the basin, and is still under development. The model is intended for simulations of extreme events, and is calibrated to reproduce the extreme flood in May 2014 as no other severe flood has been observed in the basin since this event. The problem in this process is the fact that the availability of subdaily rainfall data is much lower than that of daily data.

The stochastic generation and disaggregation models are generally used in operational hydrological practice for determination of the flash flood lead time (Lardet & Obled 1994), but not for the everyday forcing of prognostic hydrological models with disaggregated observed daily rainfall. In this paper, we introduce a modelling chain where the wflow_hbv hydrological model and multivariate rainfall disaggregation tool MuDRain (Koutsoyiannis *et al.* 2003), implemented within the MuDRain software (ITIA 2001), are coupled and applied within the existing framework of the operational flash flood forecasting system. The objective is to investigate the improvement of rainfall spatial distribution as the input to the hydrologic model using the hourly data produced with MuDRain. The wflow_hbv model is forced with hourly rainfall series observed during the event in May 2014 by the automatic rain gauge network and obtained by disaggregation of daily data in order to evaluate the effects of additional information from daily rainfall at non-recording rain gauges. Therefore, the paper is focused on the first phase of the flood forecasting process, in which the model is forced with the most recent observations to provide the initial conditions of the hydrologic model state variables prior to second phase, including model runs with forecasted rainfall.

The paper is organized as follows: details on the Kolubara basin, available hydrometeorological data, and flood in May 2014 are presented in the Study area and data section. The wflow_hbv and MuDRain models, along with the setup for the simulation of the flood event, are given in the Methodology section. Impact of the rainfall disaggregation is elaborated in the Results and discussion, while the summary and concluding remarks are given in the final section.

STUDY AREA AND DATA

Study area

The Kolubara River basin is located in the western part of Serbia. The catchment area upstream of the Draževac hydrologic station at the basin outlet is 3,636 km². The basin covers altitudes from 76 up to 1,346 m a.s.l. with fairly steep headwater areas. Deciduous forests and cultivated agricultural land prevail at higher elevations, while

the lower catchment is urbanized and densely populated (Langsholt *et al.* 2013). The Kolubara River hydrologic regime has a pronounced seasonality with flood flows mainly occurring in March and April due to combined rainfall and snowmelt, and low flows from August to October (Todorović & Plavšić 2016). Floods in the Kolubara catchment can also occur in summer due to intensive convective storms that trigger quick response of mainly torrential headwater tributaries. The Kolubara catchment has a specific fan-type shape (Figure 1), causing simultaneous occurrence of flood waves from the tributaries in the main river.

Measurements at 13 hydrologic stations within the basin (see Figure 1 and Table 1) with hourly discharge data are used in this study. Nine hydrologic stations measure runoff from drainage areas with mean elevation ranging from 150 to 450 m a.s.l., while the remaining stations measure runoff from drainage areas from 450 up to 850 m a.s.l.

Daily rainfall depths are observed at 23 non-recording rain gauges (RG) within the Kolubara basin and its vicinity (Figure 1 and Table 2). The automatic (recording) rain gauges (ARG) that measure subdaily rainfall are located at five locations next to five RGs. We shall refer to these locations as twin rain gauges. The ARGs are equipped with temperature sensors. Subdaily ARG data at two stations within the catchment are available from April 2014 and from three stations outside the catchment from 2010. RGs have much longer records (see the last column of Table 2). Over 70% of the stations are located between 50 and 250 m a.s.l., six stations (26%) cover elevations from 250 to 500 m a.s.l., and only one station is situated at 943 m a.s.l.

Flood event in May 2014

Massive flooding affected the Kolubara catchment in May 2014. The flood was induced by heavy prolonged rainfall, nearly saturated soil due to previous rainfall events and backwater effect from the high water level in the Sava River (Plavšić *et al.* 2014). The rainfall event lasted from May 12th to May 19th with the highest rainfall depths observed between 13th and 16th May. Extreme rainfall amounts were a consequence of the cyclonic air circulation generated in Genoa bay that moved and intensified



Figure 1 | The Kolubara River basin, locations of automatic rain gauges (filled circles), rain gauges (empty circles) and hydrologic stations (triangles).

Table 1 | Hydrologic stations in the Kolubara basin

Hydrologic station	River	Drainage area (km ²)	Mean elevation of drainage area (m a.s.l)	Flow observations	Previous maximum flow (m ³ /s)	Maximum flow in May 2014 (m ³ /s)
Belo Polje	Obnica	185	412	Hourly	151	247
Sedlare	Jablanica	140	808	Hourly	163	138
Valjevo	Kolubara	340	494	Hourly	287	396
Deguric	Gradac	159	665	Hourly	185	163
Mionica	Ribnica	108	557	Hourly	418	188
Slovak	Kolubara	995	430	Hourly	322	1,100
Bogovađa	Ljig	679	306	Hourly	269	195
Beli Brod	Kolubara	1,896	366	Hourly	767	1,360
Zeoke	Peštan	125	251	Daily	116	116
Koceljjeva	Tamnava	208	237	Hourly	126	178
Ćemanov Most	Tamnava	386	167	Hourly	61.2	147
Ub	Ub	214	228	Hourly	94.5	146
Draževac	Kolubara	3,588	275	Hourly	646	1,260

over the south Adriatic Sea and then persisted over the central Balkan Peninsula for several days (Nišavić et al. 2014). The following rainfall depths were observed at RGs between 10th and 19th May 2014: 328.5 mm at

Majinović, 309.3 mm at Stepojevac and 297.5 mm at Stubline. The observed rainfall at remaining RGs did not exceed 170 mm. Figure 2 presents mass curves of rainfall measured at ARGs. Accumulated rainfall depths

Table 2 | Meteorological stations in the Kolubara basin (ARG – automatic, RG – daily) and data availability (P – precipitation, T – temperature)

Rain gauge	Type	Available data		Elevation (m a.s.l)	Data available from
		Daily	Hourly		
Majinović	RG	P		303	1958 April 2014
	ARG		P, T		
Sopot	RG	P		170	1958 2010
	ARG		P, T		
Štavica	RG	P		220	1958 April 2014
	ARG		P, T		
Velika Ivanča	RG	P		225	1952 2010
	ARG		P, T		
GMS Valjevo	RG	P		185	1958 2010
	ARG		P, T		
RCValjevo	RG	P		388	1973
Lukavac	RG	P		247	1958
Brežde	RG	P		318	1958
Mionica	RG	P		170	1958
Liplje	RG	P		246	2014
Rogačica	RG	P		254	1958
Jagodići	RG	P		943	1958
Stubline	RG	P		87	1958
Stepojevac	RG	P		115	1958
Sibnica	RG	P		227	1958
Šarbane	RG	P		121	1958
Koceljewa	RG	P		131	1958
Pambukovica	RG	P		171	1958
Ub	RG	P		100	1958
Bogatić	RG	P		442	2006
Osečenica	RG	P		468	2006
Prkosava	RG	P		231	1958
Donje Crniljevo	RG	P		181	1958

at ARGs are compared to twin RGs in Table 3. Differences in rainfall amounts at ARGs and RGs are considerable at some twin gauges. Such differences are often present in operational hydrological practice (Ivković & Nađ 2015).

During the flood event there were gaps in hydrologic observations at some stream gauges due to flow recorder failures, while some instruments were washed away (Prohaska & Zlatanović 2015). The observed peak flows in May 2014 and previous maxima are compared in the

last two columns of Table 1. Previous maxima were exceeded at most stations and indicate the severity of this flood event.

METHODOLOGY

Overview of the wflow_hbv hydrologic model

The wflow_hbv is a raster-based fully distributed hydrologic model grounded in the original HBV-96 model (Bergström 1992; Bergström & Graham 1998). The model is coded in Python using PCRaster library (Karszenberg *et al.* 2010) and implemented with other supporting Python libraries (Scipy and Numpy). The PCRaster system is developed for both general-purpose raster data manipulation and as a spatially distributed dynamic modelling system. It incorporates mathematical analysis language that enables easy manipulation and analysis of spatial data.

The model comprises the precipitation, soil moisture and runoff response routines. The conceptual structure of the model is presented in Figure 3. Precipitation is treated as rainfall if the air temperature is above the defined threshold. Otherwise, it is considered snow. Snowfall is added to the dry component of the snowpack, while rainfall is directed to the free water reservoir that represents either rain or the liquid phase of the snowpack. Two components in the snow pack interact through processes of melting and refreezing, depending on the specified temperature thresholds.

Canopy interception is simulated with a single reservoir with constant capacity, which is a model parameter that can be inferred from the vegetation type. Intercepted water either constitutes throughfall or evaporates.

Snowmelt and throughfall become available for infiltration into the soil zone. Depending on the soil moisture deficit, water infiltrates until full saturation while the excess water is available for direct runoff. The seepage from the unsaturated zone is a non-linear function of the relative soil moisture content. Evapotranspiration from the soil zone also depends on the current soil moisture content. It increases linearly with the soil moisture content up to the specified threshold at which it becomes equal to potential evapotranspiration.

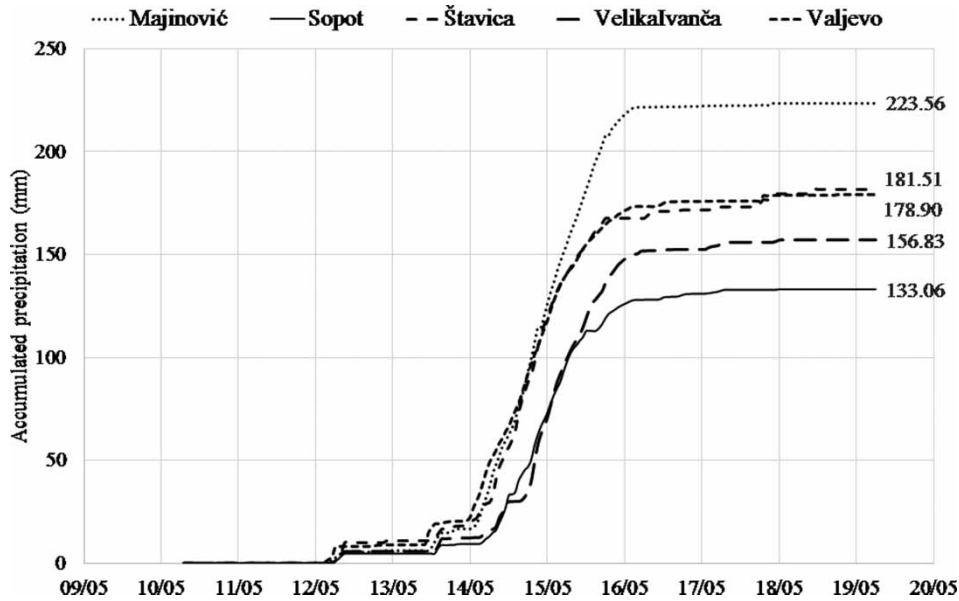


Figure 2 | Mass curves of rainfall at five ARGs from 10th to 19th May 2014.

Table 3 | Total rainfall at five ARGs and twin RGs between 10th and 19th May 2014

	Valjevo		Velika Ivanča		Sopot		Štavica		Majinović	
	ARG	RG	ARG	RG	ARG	RG	ARG	RG	ARG	RG
Total rainfall (mm)	178.9	205.6	156.9	133.3	133.1	190.5	181.5	155.9	223.6	328.5

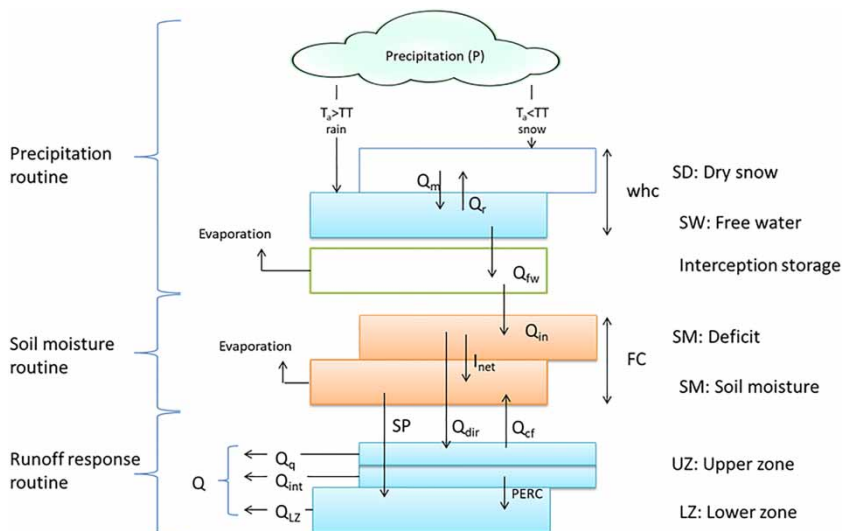


Figure 3 | Scheme of the wflow_hbv model.

Seepage water and surface runoff enter the runoff response routine, where two reservoirs simulate different runoff processes. The upper zone non-linear reservoir generates interflow and also, in the case of excess water, quick runoff. The lower zone linear reservoir generates slow runoff. In the upper zone, the processes of percolation toward the lower zone and capillary flow are also taken into account in the water balance equations. Total runoff is the sum of the interflow, quick and slow runoff.

Unlike the original HBV model, in the wflow_hbv, flow routing in a stream network is computed by using the kinematic wave method.

A total of 21 free model parameters control all hydrological processes in each grid cell. The outputs from the wflow_hbv are the following spatially distributed variables: dry and wet snow fraction, canopy storage, soil moisture content, percolation, and capillary rise lower zone and upper zone reservoir and storage, surface runoff, base flow, and total flow in a stream network.

Input data for hydrological simulations are precipitation depths, temperatures and potential evapotranspiration rates. Input data can be imported either as time series or as gridded maps. If point observations are used, spatial interpolation in wflow_hbv is performed with the inverse distance weighting method. A digital terrain model (DTM) of a basin is required, while land-use, soil or vegetation maps are desirable for parameter estimation.

Multivariate rainfall disaggregation framework

Koutsoyiannis *et al.* (2003) developed a simple and parsimonious multivariate rainfall disaggregation model that basically comprises two steps: generation of subdaily values without reference to daily totals and application of the adjustment procedure to preserve the observed daily totals. The model is implemented in the MuDRain software (available from NTUA/ITIA site; ITIA 2001) and assumes that subdaily rainfall can be well represented by an AR(1) process. The multivariate AR(1) model is given by:

$$\mathbf{X}_t = \mathbf{a}\mathbf{X}_{t-1} + \mathbf{b}\mathbf{V}_t \quad (1)$$

where $\mathbf{X}_t = [X_t^1, X_t^2, \dots, X_t^n]^T$ represents the $n \times 1$ vector of subdaily rainfall at time t ($t = 1, 2, \dots, k$), where k is the

number of subdaily periods in a day; $k = 24$ for hourly data at n locations, \mathbf{a} and \mathbf{b} are the $n \times n$ matrices of coefficients, and \mathbf{V}_t is an independent (both spatially and temporally) identically distributed vector of size $n \times 1$. The model parameters (coefficients \mathbf{a} and \mathbf{b} and moments of \mathbf{V}_t) are estimated from the statistics that should be preserved in the generated subdaily series, including: (1) the means, variances and coefficients of skewness; (2) autocorrelation function; (3) lag zero cross-correlation; and (4) proportion of dry intervals. Therefore, by identifying the AR(1) model in Equation (1), the target cross-correlation among the locations is introduced in the procedure. If subdaily data are available at p out of n locations, the first p elements of \mathbf{X}_t represent the observed rainfall and the first p elements of \mathbf{V}_t are calculated directly from the data. The rainfall series at remaining $n - p$ locations are generated.

The adjustment procedure in the model corrects the generated subdaily values in such a way that their daily totals over the whole period of disaggregation become equal to the observed daily totals without affecting the first and second order properties of the stochastic process. The set of generated subdaily vectors $\tilde{\mathbf{X}}_t$ that satisfies the accepted limit for the differences between the observed and simulated daily totals for all locations is chosen for further use. Finally, the chosen set $\tilde{\mathbf{X}}_t$ is adjusted to obtain the final generated subdaily vectors \mathbf{X} such that $\sum_{t=1}^k \mathbf{X}_t$ is equal to corresponding daily totals.

For the application of the MuDRain model, two rainfall data sets need to be prepared: daily rainfall data from rain gauges and subdaily rainfall data from automatic rain gauges. The subdaily data are used for identification of the model in Equation (1), while the daily data are actually being disaggregated. In this study, we disaggregated daily data into hourly values. Target cross-correlations between hourly rainfall data are also required for the model run. For stations with daily data the hourly cross-correlations can be estimated using an empirical relationship recommended by Koutsoyiannis *et al.* (2003):

$$r_{ij,h} = (r_{ij,d})^m \quad (2)$$

where $r_{ij,h}$ and $r_{ij,d}$ are the hourly and daily cross-correlation coefficients between stations i and j , respectively, and m is a parameter that can be assessed based on the data from twin

ARG and RG stations. Alternatively, in the case of insufficient subdaily data, the value of m can be assumed approximately in the range from 2 to 3 (Fytilas 2002).

Modelling setup

The hydrologic modelling with wflow_hbv requires a number of pre-processing activities, mainly for preparation of the static maps (terrain and other thematic maps). The digital terrain map of the basin is prepared using maps from 3 arc-seconds SRTM digital elevation model (Farr *et al.* 2007). Vegetation and land-use types are identified from the Corine Land Cover 2000 (Heymann *et al.* 1994). The soil types are obtained from European Soil Data Centre (ESDAC) (Panagos *et al.* 2012). Spatial resolution of all maps is 1 km × 1 km.

Generally, if a model is intended for simulations of such extreme floods, it should be calibrated against extreme events (Cloke & Pappenberger 2009). No extreme floods occurred after the flood in May 2014 and therefore this event is selected for calibration. The wflow_hbv model parameters are estimated through the process of calibration against the observed hourly discharge at hydrologic stations in the basin (Table 1) and with hourly precipitation depths and temperatures from five ARGs (Table 2) during the event from May 2014. All available temperature and rainfall data were included in this process because of the short data record.

The wflow_hbv model is calibrated by optimizing model parameters with NSGA-II genetic algorithm (Deb *et al.* 2002) from the AMALGAM suite of optimization algorithms (Vrugt 2016). Seven parameters are optimized using the Nash–Sutcliffe efficiency coefficient as the objective function, while the remaining parameters are set according to the basin characteristics. Six model parameters are spatially distributed. For the purpose of more efficient estimation of the distributed parameters, the grid cells are grouped for all combinations of soil types and land-use classes. The optimized model parameters are given in Table S1 in the online supplementary material.

To estimate the value of the additional hourly rainfall data on flood flow simulation, the wflow_hbv model is applied with two different inputs:

- (1) observed hourly rainfall at five ARGs;
- (2) disaggregated hourly rainfall at 23 locations in the basin.

Both wflow_hbv simulations are performed with the same wflow_hbv model setup, i.e., with the same parameter set. The same parameter set is used for simulations with both inputs in order to avoid compensation of rainfall impact by the hydrologic model parameters. The model parameters are generally affected by rainfall spatial distribution, as shown by, e.g., Chaubey *et al.* (1999), and model calibration with two rainfall data would inevitably lead to different parameter estimates. By using the same model parameters, the differences in the simulated hydrographs would be due to rainfall input only, and any increase in fit to the observations can be attributed to improved rainfall spatial representation via disaggregation. Application of the same parameter set can be also justified by the saturated soil over this period, meaning that the catchment response is primarily affected by rainfall variability, while the impact of the parameters is insignificant (Anquetin *et al.* 2010).

For the second rainfall input, the MuDRain model is used to create hourly rainfall data at 23 locations. Daily rainfall data from 23 RGs and hourly data from five ARGs between 10th and 19th May 2014 are used as the input for MuDRain. Another input for MuDRain is the target cross-correlation matrix for hourly data. Regression analysis of corresponding daily and hourly cross-correlations does not produce a meaningful estimation of the MuDRain parameter m (Equation (2)), probably due to only two months of available hourly observations (see Table S2 and Figure S1 in the online supplementary material). Therefore, instead of inferring the parameter m from the observed data, we chose to adopt a value of m according to recommendations by Fytilas (2002). A brief analysis of sensitivity of MuDRain to m in the range between 2 and 3 has shown that the resulting hyetographs have negligible differences for different values of m and we adopted $m = 2$. Using Equation (2) with daily cross-correlations calculated from long records at 23 RGs and with the adopted value of m , target hourly cross-correlations are defined as the input for MuDRain. The remaining MuDRain parameters are used with their default values (no log or power transform, zero threshold for precipitation, and only one repetition).

The output from the MuDRain are hourly rainfall series at 23 locations that are further spatially interpolated to obtain gridded input for the distributed wflow_hbv model.

Two rainfall inputs are compared in terms of hyetographs at the ARGs, mean basin precipitation depth, and spatial rainfall distribution obtained with the inverse distance weighting interpolation method.

RESULTS AND DISCUSSION

As a stochastic tool, MuDRain produces different temporal distributions of hourly rainfall at locations of RGs in every run while preserving their daily totals. At locations of ARGs, the hourly data resulting from MuDRain are not stochastic because they represent observed hourly rainfall at these sites, but adjusted to the daily totals observed at twin RGs. Figure 4 shows the results of 50 simulations from MuDRain for two selected ARG and RG locations. The left graph of Figure 4 represents the simulated hyetographs at one ARG, showing that each simulation produces the same hyetograph. The right graph in Figure 4 represents the results of MuDRain simulations at one RG, with the mean values and the 95% confidence interval. The width of the 95% confidence interval is narrow, indicating that the random component of the stochastic generation is small and consequently that variation in rainfall temporal distribution is also small. For this reason, and because the rainfall temporal distribution is the primary driver of the hydrograph shape, we can assume that running the operational hydrometeorological system with just one MuDRain simulation is reasonable if the rainfall distribution over the basin is relatively uniform (which is the case with this event).

To demonstrate the usefulness of the disaggregation method for flood simulations, simulated hydrographs with both rainfall inputs are compared in terms of input rainfall volumes and spatial distribution, and hydrograph features such as the peak flow, flow volume, start of direct runoff and time to peak.

Two rainfall inputs (ARGs only and disaggregated rainfall) are compared in terms of mean basin precipitation depth and spatial rainfall distribution obtained with the inverse distance weighting interpolation method. Mean rainfall depth in the basin for the period between 11th and 16th May is shown in Figure 5.

Spatial distributions of the differences between two rainfall inputs to hydrological model are presented in Figure S2 in the online supplementary material, while Figure 6 only shows areas of over- and underestimation between two rainfall inputs. The disaggregated rainfall depths mainly exceed the ARG observations, but vary in both space and time without exhibiting any apparent pattern. The frequency distributions of the differences between two rainfall inputs at each pixel within the basin are shown in Figure S3 in the online supplementary material.

Simulated hydrographs at selected stream gauges with the ARG observations and disaggregated rainfall are presented in Figure 7, along with the observed flows. The results for all hydrologic stations are shown in Figure S4 in the online supplementary material. These hydrologic stations are selected because only moderate flooding occurred upstream of them and the observed hydrographs are considered appropriate for comparison with hydrologic simulations. The simulated hydrographs are generally

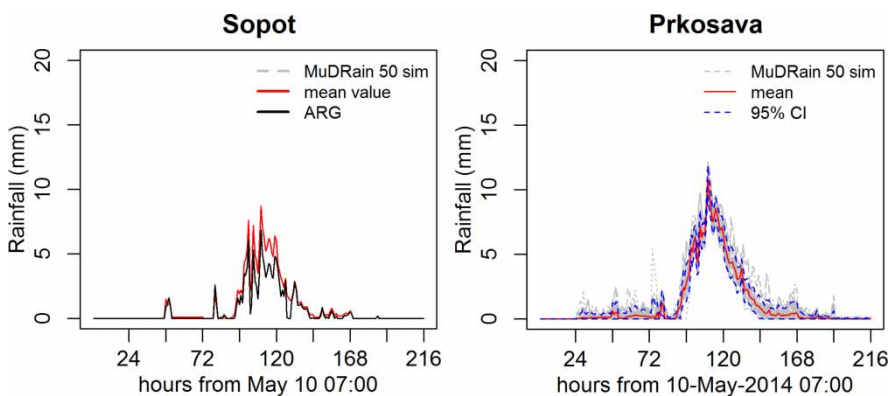


Figure 4 | Results of 50 MuDRain simulations: at Sopot ARG with mean value and the observed data (left) and at Prkosava RG (right) with mean values and 95% confidence interval.

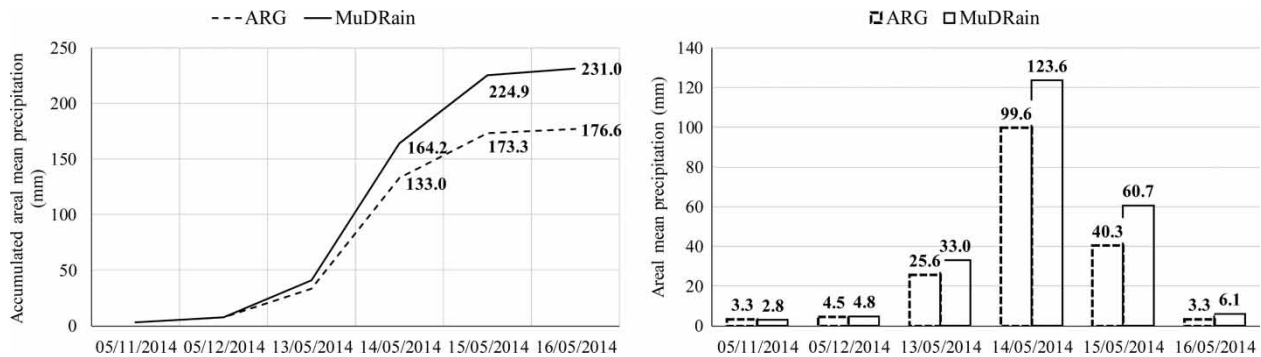


Figure 5 | Mean areal precipitation in the Kolubara basin inferred from five ARG observations and disaggregated rainfall with the MuDRain representing two rainfall inputs for the hydrologic model: mass curves (left) and daily rain depths (right).

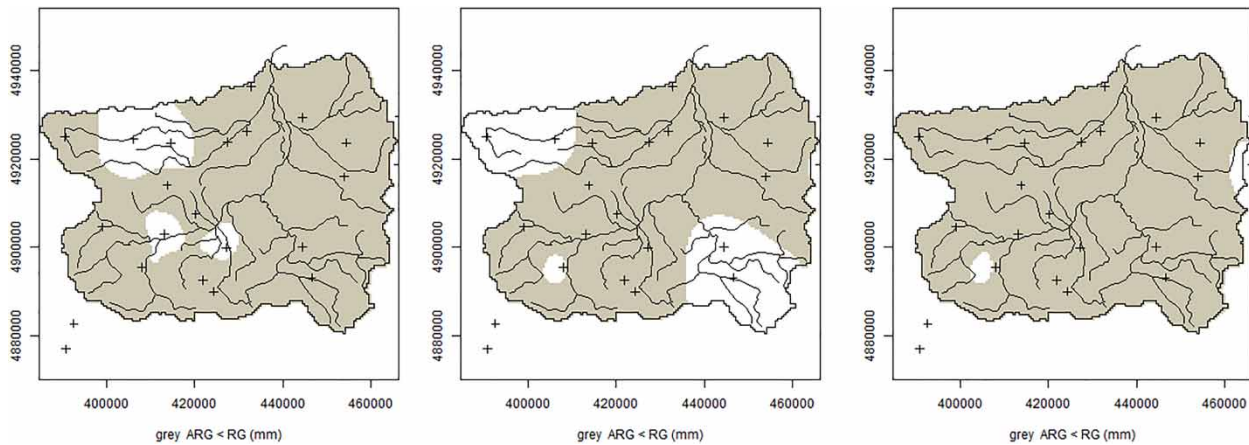


Figure 6 | Difference between daily rainfall at ARGs and from MuDRain from 13th to 15th May 2014; disaggregated rainfall exceeds ARG observations within the shaded areas and opposite within the white areas.

consistent with the observations. Considering that the model cannot simulate either flooding (water retained outside the main river channel for a longer time than the simulation period) or effect of the reservoirs, flow volume is overestimated at the majority of stations, as well as the peak discharges (Table 4). Concerning flood dynamics, relatively high values of the Nash–Sutcliffe efficiency coefficient are obtained (0.69 at Valjevo), indicating the model's satisfactory performance. Therefore, the model can be used for further analysis of rainfall disaggregation impact in flood simulation.

In general, better agreement with the observed flows is obtained with the disaggregated rainfall data. In addition, the flood volume and the maximum discharge at almost all hydrological stations are greater when disaggregated data are used (Table 4) and this can be attributed to higher rain

depth in almost the entire catchment (as illustrated in Figure 6). Lower peaks with the disaggregated data are obtained at, e.g., Bogovadja, Koceljeva and Ćemanov Most, and this is because MuDRain results in lower rain depths in these subcatchments, especially on 14th May (see mid panel in Figure 6).

The disaggregated data significantly improved model performance at the Valjevo gauge in terms of flood volume and peak discharge (top diagrams in Figure 7). More information about spatial rainfall distribution in disaggregated data invoked more precise hydrographs in the head parts of the basin. For example, at Mionica (middle diagrams in Figure 7) the rising limb and the final part of the recession are better modelled with disaggregated data, although for both rainfall inputs the flood volume is overestimated.

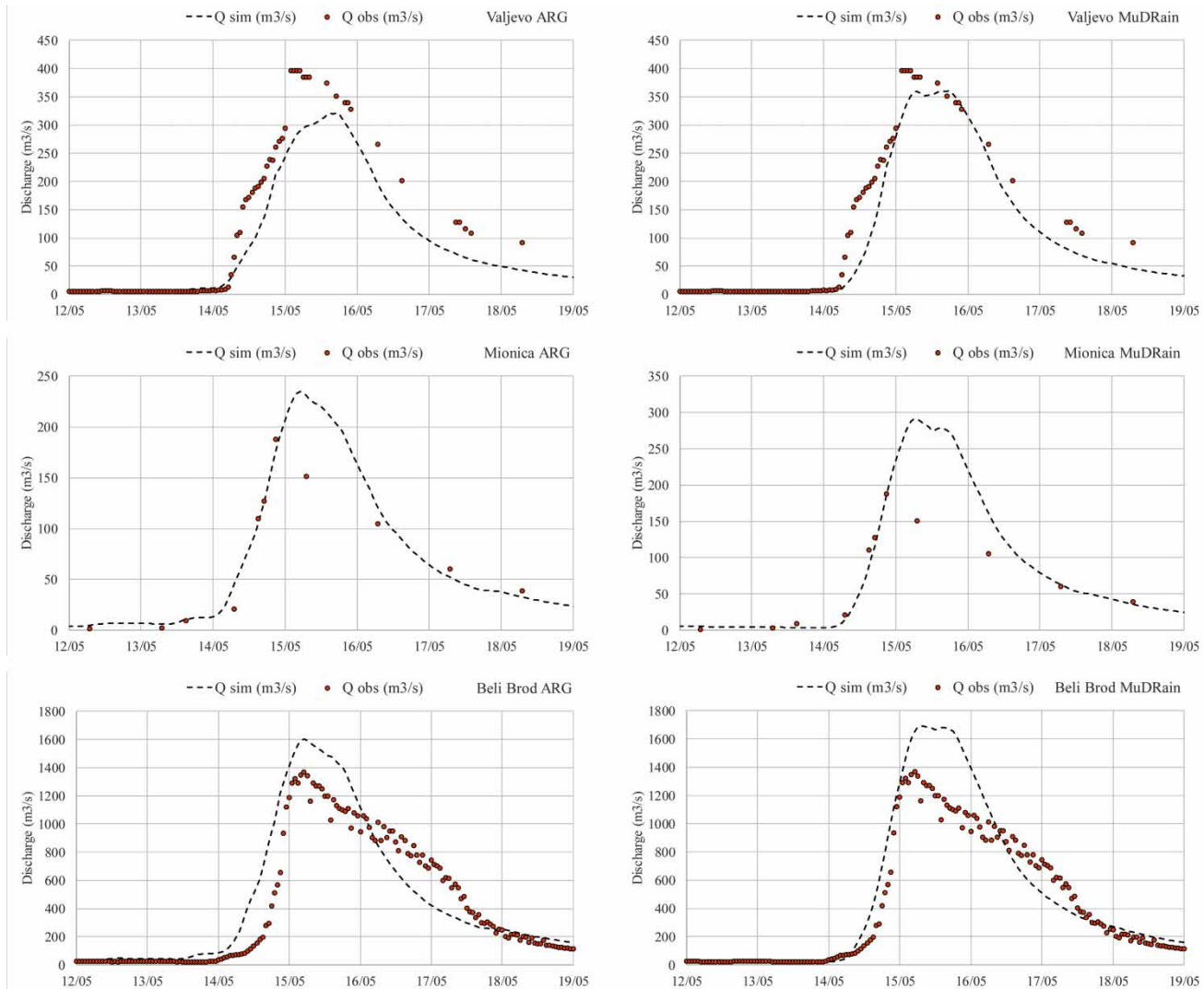


Figure 7 | Simulated (dashed line) vs. observed (dots) hydrographs at Valjevo (top), Mionica (middle) and Beli Brod (bottom). Simulations performed with the ARG observations (left) and with the disaggregated rainfall with the MuDRain model (right).

The rising limb at Beli Brod (bottom diagrams in Figure 7) is well reproduced with the disaggregated rainfall forcing, while with ARG observations the flood wave appears earlier. These results indicate that improved rainfall spatial distribution from denser rain gauge network, together with fine temporal rainfall representation, is a prerequisite for accurate simulations of flood runoff dynamics, particularly of time to peak and start of the rising limb, which are both very important for flood forecasting. At many hydrological stations, higher peak flow values and smaller bias in runoff volume are obtained with the disaggregated data (see Table 4).

CONCLUSIONS

The ability of a flood forecasting model to provide accurate forecasts is conditioned on well-estimated initial conditions and accurately forecasted precipitation and temperature. Many flood-prone areas in Serbia do not have the hourly rainfall measurements necessary for fast reaction in extreme hydrometeorological situations. It is therefore important to evaluate possible improvements in flood forecasting from readily available daily rainfall data from much denser non-recording rain gauge networks. This paper addresses this issue by performing hydrologic simulations with the

Table 4 | Characteristics of the computed and observed hydrographs

Hydrologic station	V_{obs} (10^6 m ³)	ARG V_{sim} (10^6 m ³)	MuDRain V_{sim} (10^6 m ³)	ARG V_{sim}/V_{obs} (-)	MuDRain V_{sim}/V_{obs} (-)	$Q_{max, obs}$ (m ³ /s)	ARG $Q_{max, sim}$ (m ³ /s)	MuDRain $Q_{max, sim}$ (m ³ /s)
Valjevo	84.78	59.7	67.28	0.70	0.79	396	321.2	360.97
Degurić	27.02	23.57	22.72	0.87	0.84	164	118.17	117.55
Mionica	40.45	43.88	51.53	1.08	1.27	188	234.96	290.96
Slovac	167.29	167.92	184.96	1.00	1.11	1,100	868.81	1,028.73
Bogovađa	59.83	93.72	92.87	1.57	1.55	195	553.42	493.41
Beli Brod	261.61	297.07	313.08	1.14	1.20	951	1,601.43	1,689.07
Koceljeva	28.61	37.86	27.86	1.32	0.97	178	217.5	159.0
Ćemanov Most	39.48	58.51	44.45	1.48	1.13	147	332.51	243.61
Ub	26.5	38.39	38.25	1.45	1.44	146	217.35	231.15
Zeoke	20.966	52.52	90.77	2.51	4.33	116	322.96	635.67
Draževac	477.88	552.6	629.98	1.16	1.32	1,500	2,283.5	3,634.21

operational wflow_hbv distributed model with hourly rainfall from a sparse network of recording rain gauges, and on the other hand, with hourly rainfall disaggregated from daily data from the network of non-recording rain gauges. The existing sparse recording rain gauge network in the Kolubara basin tends to underestimate rainfall during the extreme events compared to the rainfall observed at non-recording gauges. Such a situation is to be expected regularly in operational practice.

The multivariate rainfall disaggregation method proposed by Koutsoyiannis *et al.* (2003) reproduces temporal distributions at twin locations of recording and non-recording rain gauges and improves the estimates of rainfall volume. The method effectively adjusts mean areal rainfall to the measurements from the non-recording network. In the case study of the flooding event in May 2014 in the Kolubara River catchment, the greatest differences between the two rainfall inputs are noticeable in parts of the catchment with no recording rain gauges. Those amounts are by no means negligible, especially in the case of abundant flooding.

Hydrologic simulations with two types of rainfall input show an improvement in terms of consistency between the simulated and observed hydrographs when the disaggregated data are used from the denser network. The effect of disaggregated rainfall input is particularly noticeable in rising limbs of hydrographs. However, from the operational forecasting view, the results with rainfall input from the sparse recording network still provide acceptable results. This conclusion may

be limited because it is drawn from a case study with extreme rainfall volume but of moderate intensity and of fairly uniform distribution over the catchment. In the case of less extreme events, especially in the case of summer convective storms over partial catchment area, the value of disaggregated rainfall data from a denser non-recording network is expected to be greater for flood forecasting. However, in such cases of heterogeneous spatial rainfall distributions, the random component of stochastic rainfall generation may not be so small and multiple MuDRain simulations are desirable in order to produce an ensemble of rainfall spatial distributions. Further research is therefore needed to assess the usefulness of the disaggregated data from a dense network for forecasting floods from different types of storm events. Similarly, sensitivity of hydrologic simulations to the number and spatial configuration of recording gauges on which the disaggregation model is built has to be analysed in order to formulate minimum requirements for the application of this method in operational forecasting practice.

ACKNOWLEDGEMENTS

Data used are provided by Republic Hydrometeorological Service of Serbia. The research presented is partially funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, research projects TR37010 and TR37005.

REFERENCES

- Abdellatif, M., Atherton, W. & Alkhaddar, R. 2013 Application of the stochastic model for temporal rainfall disaggregation for hydrological studies in north western England. *Journal of Hydroinformatics* **15**, 555–567.
- Andreassian, V., Perrin, C., Michel, C., Usart-Sanchez, I. & Lavabre, J. 2001 Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Journal of Hydrology* **250**, 206–223.
- Anquetin, S., Braud, I., Vannier, O., Viallet, P., Boudevillain, B., Creutin, J.-D. & Manus, C. 2010 Sensitivity of the hydrological response to the variability of rainfall fields and soils for the Gard 2002 flash-flood event. *Journal of Hydrology* **394** (1–2), 134–147.
- Arnaud, P., Bouvier, C., Cisneros, L. & Domingues, R. 2002 Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology* **260** (1–4), 216–230.
- Bergström, S. 1992 *The HBV Model – Its Structure and Applications*, RH Vol. 4. SMHI, Norrköping, Sweden.
- Bergström, S. & Graham, L. P. 1998 On the scale problem in hydrological modelling. *Journal of Hydrology* **211** (1–4), 253–265.
- Blöschl, G. & Sivapalan, M. 1995 Scale issues in hydrological modelling: a review. *Hydrological Processes* **9** (3–4), 251–290.
- Chaubey, I., Haan, C., Grunwald, S. & Salisbury, J.M. 1999 Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology* **220**, 48–61.
- Cloke, H. L. & Pappenberger, F. 2009 Ensemble flood forecasting: a review. *Journal of Hydrology* **375** (3–4), 613–626.
- Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. 2002 A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6** (2), 182–197.
- Dolšak, D., Bezak, N. & Šraj, M. 2016 Temporal characteristics of rainfall events under three climate types in Slovenia. *Journal of Hydrology* **541**, 1395–1405.
- Einfalt, T., Arnbjerg-Nielsen, K., Golz, C., Jensen, N.-E., Quirnbach, M., Vaes, G. & Vieux, B. 2004 Towards a roadmap for use of radar rainfall data in urban drainage. *Journal of Hydrology* **299**, 186–202.
- Farr, T., Rosen, P., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E. & Roth, L. 2007 The shuttle radar topography mission. *Reviews of Geophysics* **45** (2), 1–33.
- Fytilas, P. 2002 *Multivariate Rainfall Disaggregation at a Fine Time Scale*. Diploma Thesis, Università di Roma 'La Sapienza', Rome. Available at: <https://www.itia.ntua.gr/en/docinfo/560/> (accessed 31 March 2017).
- Girons Lopez, M. & Seibert, J. 2016 Influence of hydro-meteorological data spatial aggregation on streamflow modelling. *Journal of Hydrology* **541**, 1212–1220.
- Guardiola-Albert, C., Rivero-Honegger, C., Monjo, R., Díez-Herrero, A., Yagüe, C., Bodoque, J. M. & Tapiador, F. 2016 Automated convective and stratiform precipitation estimation in a small mountainous catchment using X-band radar data in Central Spain. *Journal of Hydroinformatics* **19** (2), 315–330.
- Gutierrez-Magness, A. L. & McCuen, R. H. 2004 Accuracy evaluation of rainfall disaggregation methods. *Journal of Hydrologic Engineering* **9** (2), 71–78.
- Heymann, Y., Steenmans, Ch., Croissille, G. & Bossard, M. 1994 *CORINE Land Cover. Technical Guide*. Official Publications of the European Communities, pp. 1–94.
- ITIA 2001 *MuDRain – A Computer Program for Multivariate Disaggregation of Rainfall*. National Technical University of Athens, Faculty of Civil Engineering, Department of Water Resources and Environmental Engineering. Available at: <https://www.itia.ntua.gr/en/softinfo/1/> (accessed 31 March 2017).
- Ivković, M. & Nad, J. 2015 O dostupnosti i pouzdanosti podataka pri primeni hidroloških modela u realnom vremenu (On data availability and reliability for application of hydrologic models in real time). *Vodoprivreda* **47** (4–6), 350–519.
- Karssenber, D., Schmitz, O., Salomon, P. & De Jong, K. 2010 A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling & Software* **25** (4), 489–502.
- Koutsoyiannis, D., Onof, C. & Wheeler, H. S. 2003 Multivariate rainfall disaggregation at a fine timescale. *Water Resources Research* **39** (7). doi: 10.1029/2002WR001600.
- Kretzschmar, A., Tych, W., Chappell, N. & Beven, K. 2016 What really happens at the end of the rainbow? Paying the price for reducing uncertainty (using reverse hydrology models). *Procedia Engineering* **154**, 1333–1340.
- Langsholt, E., Lawrence, D., Wong, W. K., Andjelic, M., Ivkovic, M. & Vujadinovic, M. 2013 *Effects of Climate Change in the Kolubara and Toplica Catchments, Serbia*, Report No. 62, Norwegian Water Resources and Energy Directorate, Oslo, Norway.
- Lardet, P. & Obled, C. 1994 Real-time flood forecasting using a stochastic rainfall generator. *Journal of Hydrology* **162** (3–4), 391–408.
- Maksimović, Č., Bužek, L. & Petrović, J. 1991 Corrections of rainfall data obtained by tipping bucket rain gauge. *Atmospheric Research* **27**, 45–53.
- Nišavić, A., Zarić, M., Gulan, M. & Dekić, Lj. 2014 *Meteorološki uslovi u maju 2014. godine i mogućnost prognoziranja obilnih padavina (Meteorological Situation in May 2014 and Possibilities for Heavy Precipitation Forecasting)*. Republic Hydrometeorological Service of Serbia, Report. Available at: http://www.hidmet.gov.rs/podaci/dokumenti_ciril/RHMZ_Sava_Centar_referat.pdf (accessed 31 March 2017).
- Obled, C., Wendling, J. & Beven, K. 1994 The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology* **159** (1–4), 305–333.
- Panagos, P., Van Liedekerke, M., Jones, A. & Montanarella, L. 2012 European soil data centre: response to European policy support and public data requirements. *Land Use Policy* **29** (2), 329–338.

- Pappenberger, F., Beven, K., Hunter, N., Bates, P., Gouweleeuw, B., Thielen, J. & de Roo, A. 2005 Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrology and Earth System Sciences* **9** (4), 381–393.
- Pechlivanidis, I. G., McIntyre, N. & Wheeler, H. S. 2017 The significance of spatial variability of rainfall on simulated runoff: an evaluation based on the Upper Lee catchment, UK. *Hydrology Research* **48** (4), 1118–1130. doi:10.2166/nh.2016.038.
- Plavšić, J., Vladiković, D. & Despotović, J. 2014 Floods in the Sava River Basin in May 2014. In: E. Ferrari & P. Versace (eds). *Mediterranean Meeting on Monitoring, Modelling, Early Warning of Extreme Events Triggered by Heavy Rainfall*. University of Calabria, Cosenza, Italy, pp. 241–251.
- Prohaska, S. & Zlatanović, N. 2015 Hydrologic reconstruction of the flood of 2014 in the Kolubara basin – causes and consequences. *Izgradnja* **69** (11–12), 427–444.
- Sangati, M. & Borga, M. 2009 Influence of rainfall spatial resolution on flash flood modelling. *Natural Hazards and Earth System Sciences* **9** (2), 575–584.
- Schellekens, J. 2013 The wflow_hbv model, wflow 1.0RC1 documentation. Available at: http://schj.home.xs4all.nl/html/wflow_hbv.html (accessed 31 March 2017).
- Segond, M.-L., Neokleous, N., Makropoulos, C., Onof, C. & Maksimovic, C. 2007 Simulation and spatio-temporal disaggregation of multi-site rainfall data for urban drainage applications. *Hydrological Sciences Journal* **52** (5), 917–935.
- Singh, V. P. 1997 Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes* **11** (12), 1649–1669.
- Sivakumar, B. 2005 Chaos in rainfall: variability, temporal scale and zeros. *Journal of Hydroinformatics* **7** (3), 175–184.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R. & Molnar, P. 2017 Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences* **21** (3), 1359–1380.
- Todorović, A. & Plavšić, J. 2016 The role of conceptual hydrologic model calibration in climate change impact on water resources assessment. *Journal of Water and Climate Change* **7** (1), 16–28.
- Vaze, J., Post, D., Chiew, F., Rerraud, J.-M., Teng, J. & Viney, N. 2011 Conceptual rainfall-runoff model performance with different spatial rainfall unputs. *Journal of Hydrometeorology* **12** (5), 1100–1112.
- Vrugt, J. A. 2016 Multi-criteria optimization using the AMALGAM software package: Theory, concepts, and MATLAB implementation. *AMALGAM Manual*, pp. 1–69. Available at: <http://faculty.sites.uci.edu/jasper> (accessed 15 July 2017).

First received 8 April 2017; accepted in revised form 2 February 2018. Available online 6 March 2018