

Article

High-Resolution Discharge Forecasting for Snowmelt and Rainfall Mixed Events

Tomasz Berezowski *  and Andrzej Chybicki

Department of Geoinformatics, Faculty of Electronics, Telecommunication and Informatics, Gdansk University of Technology, Gabriela Narutowicza 11/12, 80-233 Gdansk, Poland; andrzej.chybicki@eti.pg.edu.pl

* Correspondence: tomberez@eti.pg.edu.pl; Tel.: +48-58-347-18-58

Received: 23 November 2017; Accepted: 3 January 2018; Published: 10 January 2018

Abstract: Discharge events induced by mixture of snowmelt and rainfall are strongly nonlinear due to consequences of rain-on-snow phenomena and snowmelt dependence on energy balance. However, they received relatively little attention, especially in high-resolution discharge forecasting. In this study, we use Random Forests models for 24 h discharge forecasting in 1 h resolution in a 105.9 km² urbanized catchment in NE Poland: Biala River. The forcing data are delivered by Weather Research and Forecasting (WRF) model in 1 h temporal and 4 × 4 km spatial resolutions. The discharge forecasting models are set in two scenarios with snowmelt and rainfall and rainfall only predictors in order to highlight the effect of snowmelt on the results (both scenarios use also pre-forecast discharge based predictors). We show that inclusion of snowmelt decrease the forecast errors for longer forecasts' lead times. Moreover, importance of discharge based predictors is higher in the rainfall only models then in the snowmelt and rainfall models. We conclude that the role of snowmelt for discharge forecasting in mixed snowmelt and rainfall environments is in accounting for nonlinear physical processes, such as initial wetting and rain on snow, which cannot be properly modelled by rainfall only.

Keywords: discharge forecasting; WRF; numerical weather forecasting; urbanized catchment; snowmelt; rainfall; random forests

1. Introduction

Discharge events induced by mixture of snowmelt and rainfall are strongly nonlinear due to consequences of rain-on-snow phenomena and snowmelt dependence on energy balance. Such events should be of major consideration especially in urban catchments, where they can promote flooding [1]. The hydrological processes related snowmelt and rainfall mixed events, mostly concerning the rain-on-snow phenomena, were simulated or analysed recently in a number of studies [2–5]. However, snowmelt and rainfall mixed events received relatively little attention in high-resolution discharge forecasting, especially in study areas where snow processes tend to be disregarded.

Discharge forecasts are often forced using numerical weather prediction (NWP) data in regional and local scale sites all over the world. Rogelis and Werner [6] used Weather Research and Forecasting (WRF) model forecasting ensembles for timely prediction of flash floods in mountain areas in tropical regions. Li et al. [7] coupled WRF with hydrological Liuxihe model to extend the flood forecast lead time. A different approach was proposed by Tao et al. [8], where an Integrated Precipitation and Hydrology Experiment (IPHEX-IOP) was performed to characterize flood predictability dedicated to complex terrains of Southern Appalachians. This research was carried out using the Coupled surface–groundwater Hydrology Model (DCHM) forced by hourly precipitation fields and other forecasts produced by the NASA-Unified WRF (NU-WRF) model. Some studies focused on the real-time applicability of the system. For instance, an operational framework induced by meteorological

forecasts for Kavango River basin in Africa [9] and a neuro-fuzzy system for flash floods warning for the area of mountainous region of Rio de Janeiro state in Brazil [10]. The numerous studies show how NWP can be used for discharge forecasts; however, most of the attention is paid to the precipitation fields that appear to be the major driver of discharge events.

On the other hand, NWP can provide far more complete representation of data, such as snow related variables, such as snow cover, depth, water equivalent, albedo or melt. These variables can be used for enhancing the modelling of hydrological processes, such as floods, with snow processes. Numerous configuration of microphysical schemes allow for looking at the snow processes with various levels of complexity. Among the most often used models available in WRF that include snow pack processes are Community Land Model (CLM) [11,12], Rapid Update Cycle (RUC) [13] or Noah [14]. According to the literature, the models vary in the performance of snow processes simulation; however, they all provide comparable, operational valid output [15,16].

WRF with the aforementioned microphysical configurations was used in cold environments: Förster et al. [17] conducted catchment-scale simulations, where snow models were integrated into the hydrological modelling system PANTA RHEI forced by WRF meteorological forecasts in the Harz Mountains, Germany, whereas, Wu et al. [18] and Zhao et al. [19] used WRF to force snowmelt runoff in snow dominated, high mountain ranges of Central Asia. The results indicated good skill of WRF derived snow related variables for hydrological modelling, yet the studies were conducted in snow dominated not urbanized environments, which do not exhibit extensive snowmelt and rainfall mixing.

Discharge response to snowmelt and rain-on-snow events in urbanized catchment is different than in other areas. Snowmelt and rainfall mixed events in urban catchments intensify discharge while urbanization increases [20]. Moreover, the onset of spring snowmelt discharge is quicker in urbanized than in non-urbanized catchments [21]. The complexity of urbanized catchments response to snow related processes is depicted by the fact that, in such catchments, higher simulation performance (expressed in Nash–Sutcliffe efficiency) can be obtained when hydrological simulations are conducted with full energy balance model in reference to the degree-day method [22]. Despite these contrasts of urbanized and non-urbanized catchments, there were no attempts to predict mixed snowmelt and rainfall discharge events in these areas using forcing data from weather forecasting models, such as WRF. While such experiments could highlight feedback between rainfall and snow processes that should be taken into account for modelling also in other areas.

The aim of this study is to highlight the effect of neglecting snowmelt in high-resolution discharge forecasts in mixed snowmelt and rainfall catchments. We conduct the study in an urbanized catchment where the influence of snow processes on discharge is enforced. Our aim is realized by simple post-processing of the WRF in order to provide input data for a machine learning algorithm that forecasts discharge and highlights the predictors' importance.

2. Methods

2.1. Discharge Forecasting

Discharge at the catchment outlet is forecasted using Random Forests, a machine learning algorithm used for classification and regression [23]. The main principle of Random Forests is to find the optimal division of predictors' hyperspace for discrimination or quantification of a predicted variable. This is achieved by growing a tree with a random subset of features and finding the best splitting point for each of the selected features. The random trees are then grown multiple times and their output is combined by voting for classification or averaging for regression. Use of Random Forests in this study was motivated by the following reasons:

1. A catchment response, expressed in discharge, to meteorological and climatological forcing is nonlinear. Hence, a nonlinear model is required to represent this system.

2. Our approach uses big models, with multiple predictors that may vary in their significance for the forecast output. Hence, a multiple model resistive to overfitting is required to regress the variables.
3. Our aim was also to gain insight into the forecasting model behaviour, for which a predictor's importance estimation feature of Random Forests can be used.

The forecasting models predict discharge (q_t^{sr} or q_t^r) [m^3s^{-1}] at the catchment outlet at forecast time $t \in \langle 1; 24 \rangle$ [h] based on catchment average snowmelt (s_l) and rainfall (r_l) fluxes [mm/h] at lag time $l \in \langle -24; t \rangle$ [h], mean discharge (qm) [m^3s^{-1}] in 25 h preceding the forecast, i.e., mean discharge between $l \geq -24$ and $l \leq 0$, and qd [m^3s^{-1}] being a difference between qm and discharge at $l = 0$ h (the last available discharge before forecast):

$$q_t^{sr} = f(s_{-24}, \dots, s_t, r_{-24}, \dots, r_t, qm, qd). \quad (1)$$

The full model (Equation (1)) is further referred to as snowmelt and rainfall model. To assess the effect of snowmelt flux inclusion in the model, we also use a forecasting scenario with rainfall and discharge predictors only (further referred to as rainfall only model):

$$q_t^r = f(r_{-24}, \dots, r_t, qm, qd). \quad (2)$$

Effectively, the number of predictors depends on the forecast time t : the snowmelt and rainfall model (Equation (1)) will consist of 54–100 predictors and the rainfall only model (Equation (2)) will consist of 28–51 predictors. This relatively big model design allows us to look at the predictors' importance and various lag time l and at various forecast time t . We assess the importance with the increase in mean squared error (IncMSE) (%) measure that is calculated by Random Forests algorithm for each predictor. IncMSE is calculated as a difference between the out-of-bag (OOB) mean squared error (MSE) from a forest with a permuted predictor and an original predictor, next, this difference is normalized to the MSE of a forest with the original predictor and expressed as a percentage. The higher the IncMSE, the more important the predictor is. The advantage of this approach is that it gives unbiased estimation of MSE and can be easily interpreted in terms of importance of each predictor for the model [23].

The training period for the model is 10 February 2013 to 30 May 2013 and the validation period is 24 October 2012 to 9 February 2013. This winter was chosen because of the longest and most fluctuating snow cover among the period with available 1 h resolution discharge data (i.e., 2010–2016) [24]. Both in training and validation period rainfall only, snowmelt only and mixed snowmelt and rainfall events are observed (Figure 1). The forecast temporal resolution is 1 h and for each time step discharge is forecasted from $t = 1$ to $t = 24$ h. The only observations used as the predictors are the preceding discharge records: qm and qd . Snowmelt and rainfall data for both positive and negative time lags (l) are derived from WRF output, i.e., they are meteorological forecasting results from the previous or forthcoming 24 h. Perhaps, in an operational model, the preceding 24 h forcing data could be obtained from interpolation of observations and only the forthcoming forcing data would be forecasted. However, in this study, our intention was to keep one source of predictors for equivocal analysis of their importance.

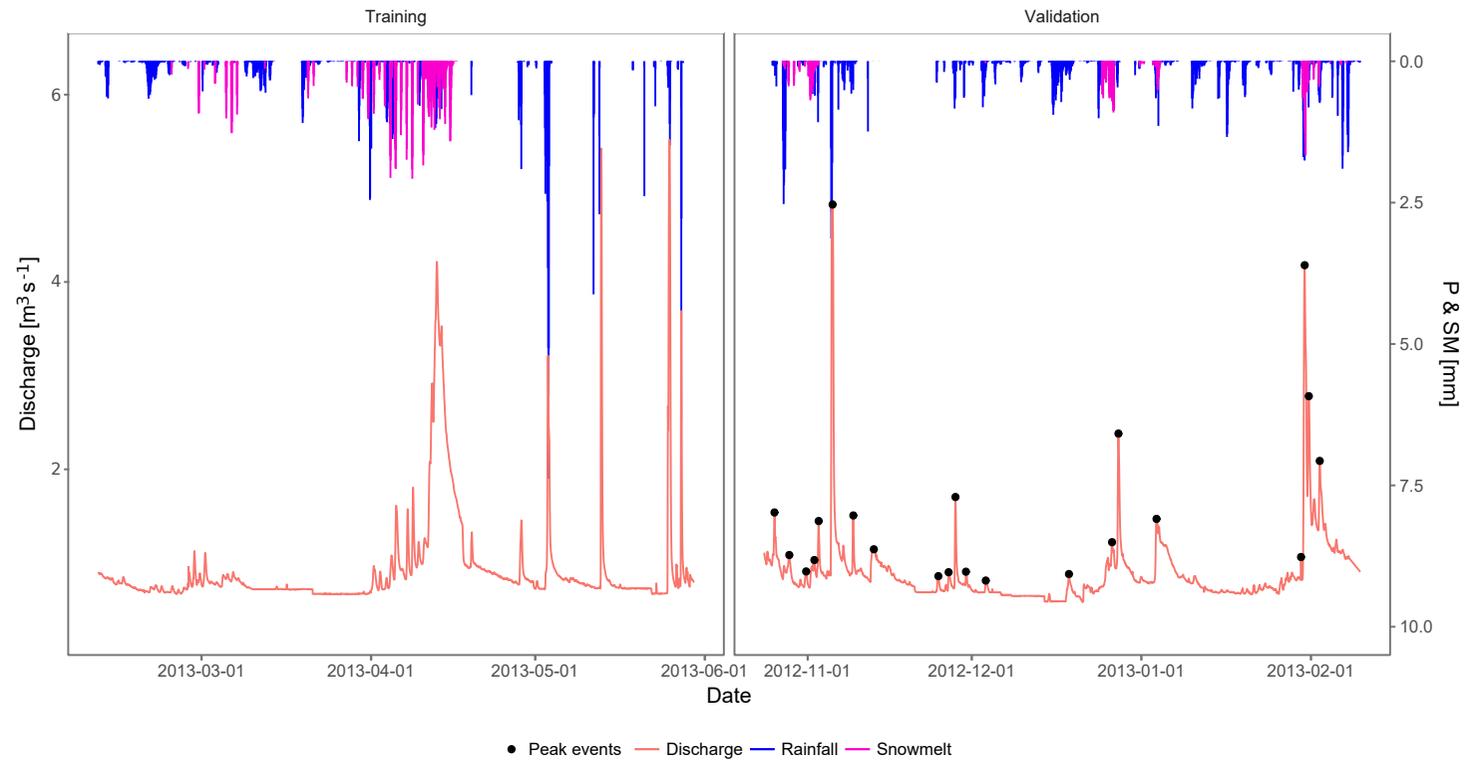


Figure 1. Training and validation time series of discharge (red), rainfall (blue) and snowmelt (pink) in 1 h resolution. The peak events used for error calculation are indicated with black dots. P stands for rainfall and SM stands for snowmelt on the right-hand side vertical axis.

The discharge forecast errors are quantified using the root mean square error [m^3s^{-1}]:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (q_i - \hat{q}_i)^2}, \quad (3)$$

and the mean error [m^3s^{-1}]:

$$\text{ME} = \frac{1}{N} \sum_{i=1}^N (q_i - \hat{q}_i), \quad (4)$$

where q_i and the \hat{q}_i are the observed and forecasted discharge [m^3s^{-1}] at time i ranging from 1 to N [h]. The RMSE quantifies the error magnitude, and the optimal RMSE is $0 \text{ m}^3\text{s}^{-1}$. The ME informs whether the observed discharge is underestimated (negative values) or overestimated (positive values), with the optimal value being $0 \text{ m}^3\text{s}^{-1}$.

2.2. The WRF Model Setup and Output

For the purpose of the study, the operational setup of WRF model instance that covers the Central Europe area was prepared. The model operates in two computational domains, namely in a parent domain with a horizontal resolution of $12 \times 12 \text{ km}$ (d01) that uniforms lateral boundary condition, and nested domain of $4 \times 4 \text{ km}$ resolution, which provides the model data used an actual input in discharge forecasting. The nested (d02) domain consists of nearly 85 thousand discrete cells (298×285) with its geographical range presented in Figure 2. The model was started every day from the Global Forecast System (GFS) reanalysis dataset producing 24 h simulations. In order to model sub-scale surface physical properties, the Noah land surface model scheme [14,15] and the Kain–Fristch cumulus scheme were used in order to account for implicit parameterization of sub-scale precipitation events. In this study, we used the WRF hourly total rainfall (r_t) and snowmelt (s_t) fields as the forecasting models (Equations (1) and (2)) predictors. The WRF model was set up to use the Ferrier operational microphysics scheme and the Mellor–Yamada–Janjic planetary boundary layer scheme [25]. In order to properly assess the convection parameterization, the utilized WRF instance provides vertical grid distribution (VGD) of 39 pressure levels over analysed computing domains with the top model layer at 5000 Pa. Proper VGD allows for explicit solving of advection schemes and better assessment of cloud microphysical processes and rainfall events. In order to account for radiation budget, model uses longwave and shortwave radiation schemes that are called every 30 min of the simulation (as default value in WRF simulations). In order to increase efficiency of the model, we used Rapid Radiative Transfer Model long wave radiation scheme [26] and simple downward integration short wave radiation scheme [27].

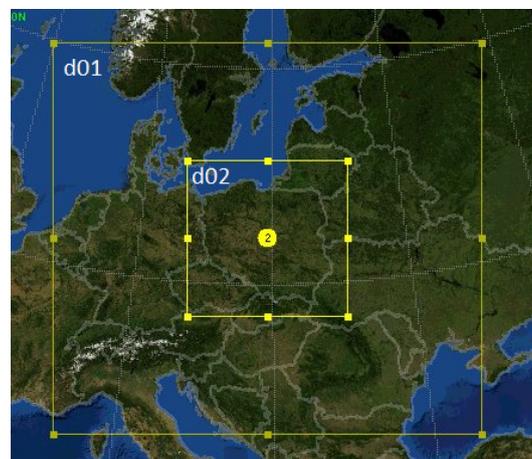


Figure 2. Geographical coverage of nested domains used in the WRF simulation.

We estimate the accuracy of WRF simulations using the observed precipitation and snow depth provided by National Oceanic and Atmospheric Administration (NOAA) for a station located in the study area [28]. One run of the 24 h forecast using the WRF model in this configuration took about 80 min; however, including the data import and export time, it increased to 90–100 min.

2.3. Experimental Setup

Our experimental setup is rather simple due to using the machine learning (“black-box”) approach as a key component (Figure 3).

The computations in our experiment were split into three computers: (1) WRF simulations were conducted on our WRF dedicated Ubuntu 10.04 workstation with four core 3.30 GHz i5 processor and 8 GB RAM; (2) WRF output grid data were post-processed into tabular time series on a high-performance computing (HPC) node with twelve core 2.3 GHz Xeon E5 processor and 256 GB RAM; (3) the random forest models for discharge predictions were trained and operated on a desktop computer using the R environment [29] and the randomForest package [30]. Note that an operational implementation of such a system would not require three computers, including an HPC, but one workstation would be enough to achieve the operational status. Our setup was motivated by the requirements of the prototyping stage in which we had to optimise our work by testing and repeating the experiment several times.

The post-processing of WRF output grid data into tabular time series was required in order to provide a proper training data-set format for the Random Forests models. In post-processing, we calculated a catchment average of WRF snowmelt and rainfall for a given hour, each of the preceding 24 h and each of the forthcoming 24 h (i.e., we calculated the s_l and r_l predictors for each l). At this step, we also included in the training data-set the qm and qd predictors, which were calculated (see Section 2.1) from the 1 h resolution discharge time series from a gauging station Zawady located at the Biala River catchment outlet (Figure 4) [24].

The Random Forests models in the regression mode were set to grow 500 trees and to calculate importance for predictors (i.e., the IncMSE measure). The OOB data was permuted once (default) while calculating importance. We did not use any predictors’ selection technique priori to Random Forests training because the importance measure effectively shows which predictors are important (and the aim of this study was to look at the importance pattern of all predictors).

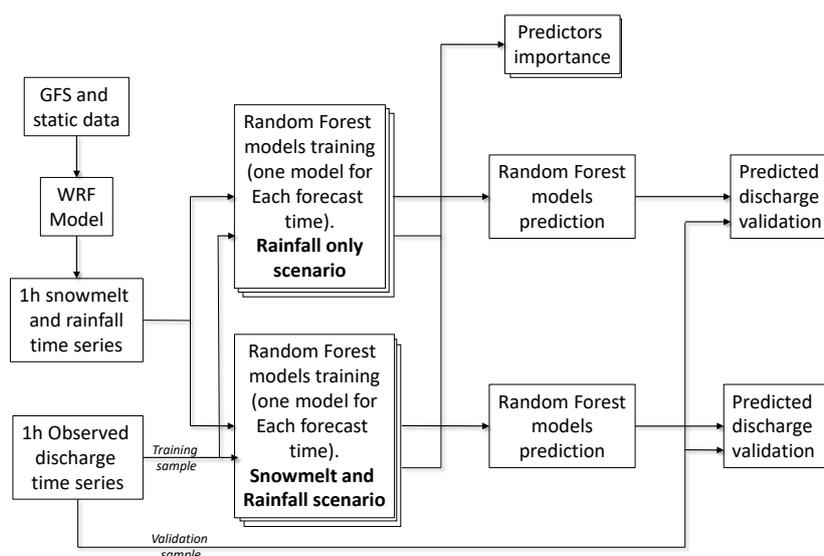


Figure 3. A flow chart presenting the experimental setup used in this study.

2.4. Study Area

Study area is the Biala River catchment 105.9 km² located in the northeastern Poland (Figure 4), where the longest persisting snow cover occurs in the lowland Poland. Climate type is humid continental according to Köppen's classification; however, the study area is located at the most western part of this type being effectively in a transitions zone between humid continental and temperate oceanic types [31]. Biala River catchment is dominated by urban land-use (46%) with minor share of agriculture (39%), forests (13%) and water (2%) [32]. Biala River flows through Bialystok City where almost each year flooding cause considerable loss. The biggest flooding is often driven by spring and summer storms, which produce high and quick discharge events; however, similar quick events also occur during winter and spring forced by mixed snowmelt and rainfall.

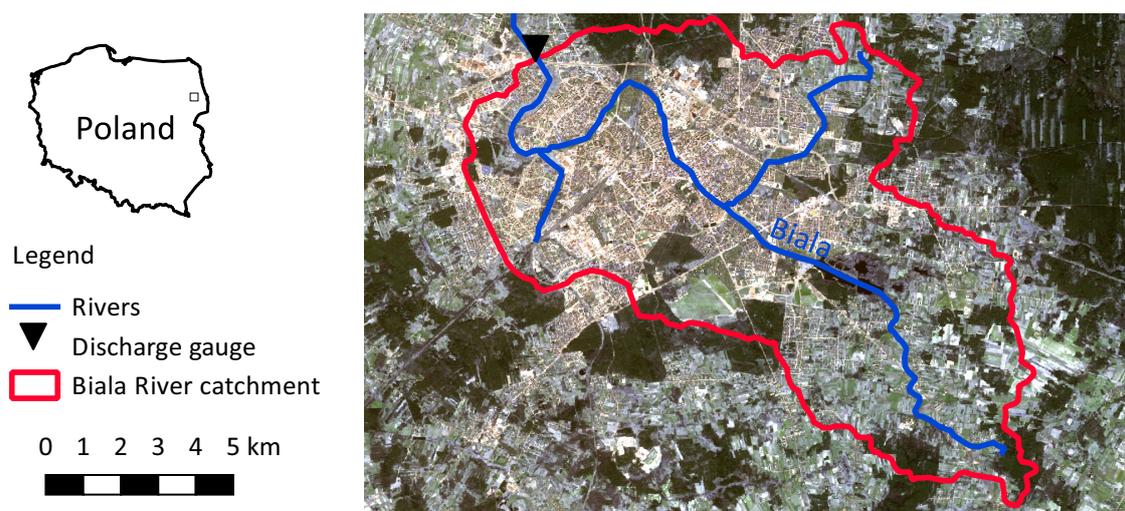


Figure 4. Biala River catchment with rivers and discharge gauge indicated. The background natural colour composition Sentinel 2 satellite image from 2017 presents urban areas in bright yellow and white, agriculture in bright green and grey and forest in dark green colour.

3. Results

3.1. WRF Results

The WRF forecasted rainfall and snowmelt matched well the meteorological observations from a station located in the central part of the catchment (Figure 5). WRF rainfall is significantly correlated with the meteorological observation ($\rho = 0.59$, $p < 2.2 \times 10^{-16}$) and has RMSE = 3.0 mm/day, which is equivalent to 13% of the observed rainfall range. Some false positive rainfall events of average magnitude equal to 0.5 mm/day can be observed in the WRF forecasts. These false positive events occur in 94 days in the training and validation period (42% of days). The false negative events are rare and occur in total in five days (2% of days). The WRF snow melt was not possible to be validated against observations because of a lack of snow water equivalent data. Nonetheless, its relation with the snow depth depletion is significant ($\rho = 0.52$, $p < 2.2 \times 10^{-16}$), which confirms a good quality of snowmelt forecasts.

The total rainfall in the calibration and validation period accounted for 269 mm and the total snowmelt for 55 mm, i.e., 20% of the total rainfall.

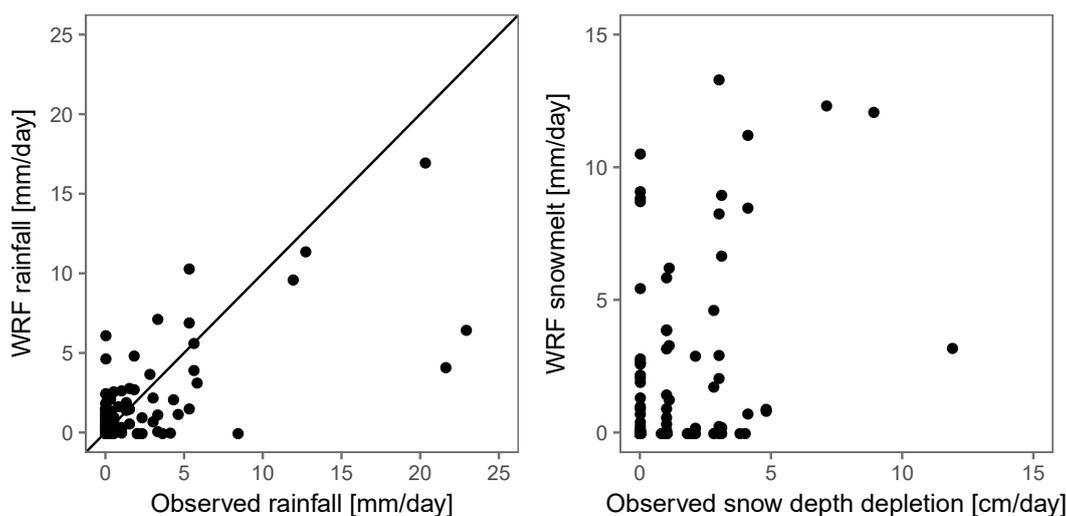


Figure 5. Validation of WRF rainfall (left panel) and snowmelt (right panel) forecasts against meteorological observations. The WRF forecasts are aggregated from hourly to daily data in order to match the meteorological records. Black line in the left panel is the 1:1 line.

3.2. Discharge Forecasting

The RMSE of the forecasted total discharge in the validation period gradually increases from $0.19 \text{ m}^3\text{s}^{-1}$ (41% of the discharge std. dev.) and $0.14 \text{ m}^3\text{s}^{-1}$ (30% of the discharge std. dev.) for snowmelt and rainfall and the rainfall only models, respectively (Figure 6). The RMSE of the snowmelt and rainfall models reach the maximum of $0.43 \text{ m}^3\text{s}^{-1}$ (92% of the discharge std. dev.) at $t = 16 \text{ h}$; however, since $t = 14 \text{ h}$, the RMSE stabilizes at around $0.42 \text{ m}^3\text{s}^{-1}$, whereas, the RMSE for the rainfall only models continue to increase until the maximum RMSE = $0.48 \text{ m}^3\text{s}^{-1}$ (102% of the discharge std. dev.) at $t = 24 \text{ h}$.

The ME of the forecasted total discharge in the validation period is at a similar level of around $0.01 \text{ m}^3\text{s}^{-1}$ with optimum ME = $2 \times 10^{-5} \text{ m}^3\text{s}^{-1}$ (0.004% of the discharge std. dev.) at $t = 24 \text{ h}$ and pessimum ME = $0.02 \text{ m}^3\text{s}^{-1}$ (4% of the discharge std. dev.) at $t = 15 \text{ h}$ for the snowmelt and rainfall models (Figure 6), whereas the ME for the rainfall only models continuously increases from optimum ME = $0.01 \text{ m}^3\text{s}^{-1}$ (2% of the discharge std. dev.) at $t = 1 \text{ h}$ until the pessimum $0.07 \text{ m}^3\text{s}^{-1}$ (15% of the discharge std. dev.) at $t = 24 \text{ h}$.

The RMSE for the forecasted peak discharge in the validation period is similar for both model variants (Figure 6). It gradually increases from $0.43 \text{ m}^3\text{s}^{-1}$ (32% of the peak discharge std. dev.) and $0.41 \text{ m}^3\text{s}^{-1}$ (31% of the peak discharge std. dev.) for snowmelt and rainfall and the rainfall only models, respectively. The peak discharge RMSE reach the maximum of $1.33 \text{ m}^3\text{s}^{-1}$ (100% of the peak discharge std. dev.) at $t = 19 \text{ h}$ for the snowmelt and rainfall models and $1.34 \text{ m}^3\text{s}^{-1}$ (101% of the peak discharge std. dev.) at $t = 14 \text{ h}$ for the rainfall only models. However, since $t = 9 \text{ h}$, both model variants have their RMSE stabilized around $1.3 \text{ m}^3\text{s}^{-1}$.

The ME for the forecasted peak discharge in the validation period is always closer to the optimal zero for the snowmelt and rainfall models than for the rainfall only models (Figure 6). The pessimal ME is $-0.61 \text{ m}^3\text{s}^{-1}$ (46% of the peak discharge std. dev.) for the snowmelt and rainfall models, which is considerably better than for the rainfall only models ME = $-0.70 \text{ m}^3\text{s}^{-1}$ (53% of the discharge std. dev.).

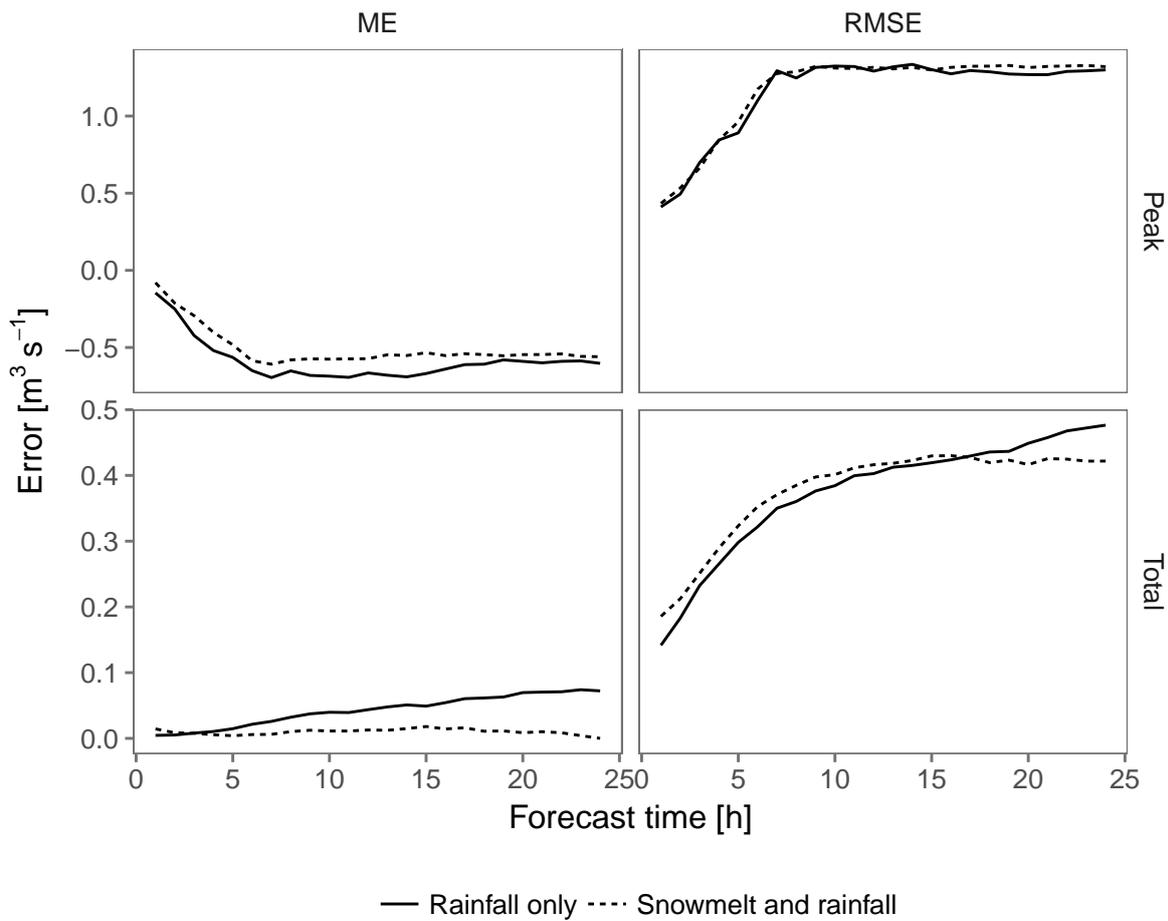


Figure 6. The mean error (ME, left panels) and root mean square error (RMSE, right panels) of the discharge forecasted with the Random Forests models for the validation period for the forecast time t from 1 to 24 h. Top panels present errors for the peak discharge events as indicated in Figure 1; the bottom panels present the total validation time series errors.

The forecasted discharge hydrograph matches the observed discharge well for the snowmelt and rainfall and rainfall only models (Figure 7). The low flows have good matching for all forecast times t . The peaks matching decrease with the increasing forecast time (cf. with Figure 6). False positive discharge peaks can be depicted for both snowmelt and rainfall and rainfall only models. However, the false positive discharge peaks appear in different situations in each model variant.

In the snowmelt and rainfall models, the false positive peaks appear during long and intensive snowmelt events, e.g., at the beginning of November or the mid-December (Figure 7), whereas, in the rainfall only models, the false positive peaks appear after snowmelt followed by rainfall events e.g., at the beginning of January or beginning of February (Figure 7). Some minor false positive discharge peaks can be observed appearing the same in both model variants during the low flow periods. The magnitude of false positive peaks increases with the increasing forecast time, and this effect is especially visible for the mixed events in the rainfall only models.

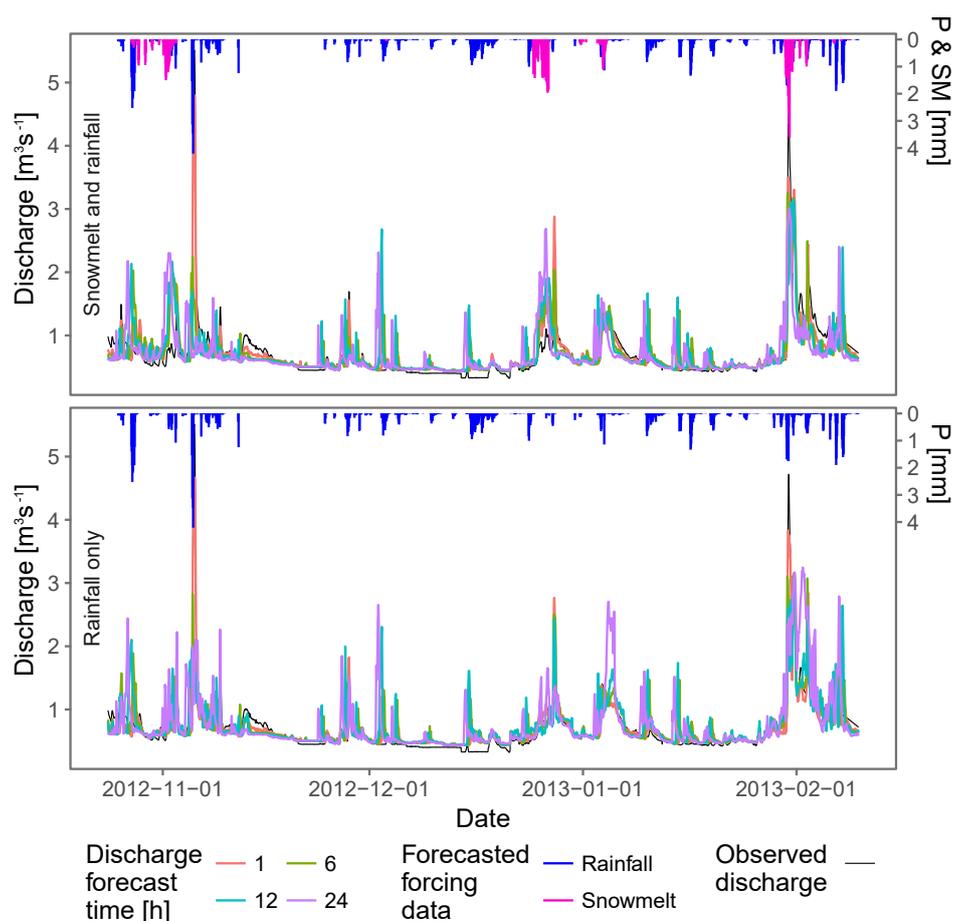


Figure 7. Forecasted discharge at four forecast times ($t = 1, 6, 12, 24$ h) for the snowmelt and rainfall models (**top** panel) and for the rainfall only models (**bottom** panel). P stands for rainfall and SM stands for snowmelt in the right hand side vertical axes.

3.3. The Forecasting Models Structure

The inclusion of snowmelt predictors in the snowmelt and rainfall models changes the predictors importance patterns in reference to the rainfall only models (Figure 8). The predictor importance pattern shows that rainfall predictors are the most important at the lag time l equal to the forecast time t . Another important rainfall predictors appear lagged 14 h to the forecast time. This is observed for both model variants. However, in the rainfall only model variants, the most lagged predictors, i.e., at $l = -24$ h, are clearly important, which is not the case in the snowmelt and rainfall models. The snowmelt predictors, opposite to the rainfall predictors, show relatively low importance at the lag time equal to the forecast time, whereas the snowmelt predictors lagging 17 h and 27 h from the forecast time have high importance. The importance of snowmelt predictors at this lag is clearly higher than the corresponding rainfall predictors, and almost as high as for the most important rainfall predictors. A clearly visible feature is that the importance of the least lagged meteorological predictors in the models forecasting at time $t = 1$ and $t = 2$ h is lower than for models forecasting at longer t .

The inclusion of snowmelt predictors in the forecast models changes the discharge-based predictors pattern in reference to the rainfall only models (Figure 9). For the snowmelt and rainfall models, the importance of discharge predictors clearly decreases with the increase of the forecast time t . On the contrary, the importance of discharge-based predictors is similar for all forecast times in the rainfall only models. The discharge difference predictor (qd) is clearly more important than the mean discharge predictor (qm) in both model variants. However, in the snowmelt and rainfall models, the importance of both predictors is more balanced than in the rainfall only models.

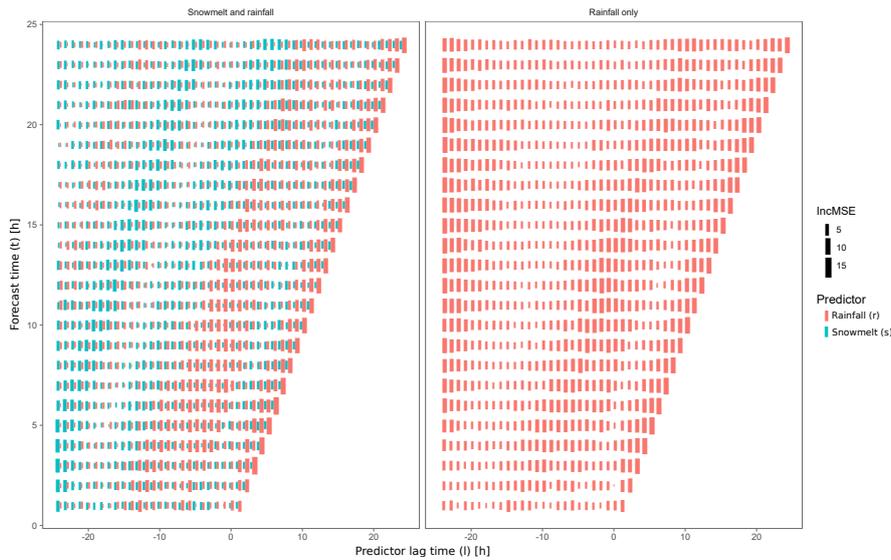


Figure 8. Importance of the forecasted meteorological predictors used in the Random Forests models (Equations (1) and (2)) for snowmelt and rainfall (**left panel**) and rainfall only (**right panel**) model variants. The vertical axis presents different forecast time $t \in \langle 1; 24 \rangle$ h; the horizontal axis presents the predictor lag time $l \in \langle -24; t \rangle$ h. The area of the rectangles is proportional to increase of the mean squared error (IncMSE) of Random Forests model after perturbing a predictor in reference to an original predictor, i.e., the higher the increase of the mean squared error, the higher the predictor importance for a model. Red rectangles present rainfall predictors (r_l) and the blue rectangles present the snowmelt predictors (s_l).

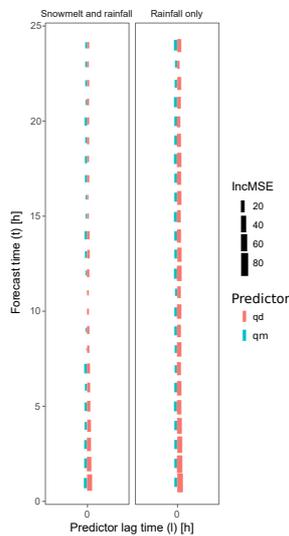


Figure 9. Importance of the discharge-based predictors used in the Random Forests models (Equations (1) and (2)) for the snowmelt and rainfall (**left panel**) and rainfall only (**right panel**) model variants. The vertical axis presents different forecast time $t \in \langle 1; 24 \rangle$ h; the horizontal axis presents the predictor lag time $l \in \langle -24; t \rangle$ h, which is always equal to 0 for this group of predictors. The area of the rectangles is proportional to the increase of the mean squared error (IncMSE) of Random Forests model after perturbing a predictor in reference to an original predictor, i.e., the higher the increase of the mean squared error the higher the predictor importance for a model. Red rectangles present the discharge difference (qd) predictor and the blue rectangles the mean discharge (qm) predictor (see Section 2.1).

4. Discussion

4.1. WRF Forecasts

The major issue concerning the WRF simulations are the false positive and false negative events observed both for snowmelt and rainfall. In our opinion, the false positive events in each case are due to the fact that WRF simulates rainfall and snowmelt fields in 4×4 km grid and the observations are recorded in a single point using a rainfall gauge. Despite the high frequency of false positive events in rainfall simulations, their magnitude is rather small and accounts for 0.5 mm/day on average. Given that the RMSE for rainfall accounts for 13% of the data range and the correlations of snowmelt and rainfall data with observations are significant and positive, in our opinion, the WRF simulation results match well the observations. Nonetheless, the false positive and negative events clearly negatively affect the Random Forests models training and forecasting. Unfortunately, such mismatch between NWP and observations is unavoidable. Several strategies exist for improving general performance of NWP, such as ensemble forecasting schemes that are aimed at finding the best WRF setup for specific cases [33,34]. However, ensemble forecasting requires huge computing power (especially for high-resolution grids and larger areas) and, importantly, the optimal setup for re-analysis is not always optimal for operational forecasting [35]. Therefore, in our case, we used a standard WRF setup that in the longer term gives the best and stable results for the analysed area [36].

4.2. Random Forests Models

In this study, we used the Random Forests model for the discharge forecasts, and this approach can be criticized due to few reasons. First, the functioning of a trained Random Forests model is difficult to understand. This is similar to neural network models, but, unlike the linear regression models, which have a very clear structure. Next, the models we used are big, i.e., from 28 to 100 predictors, which can also be understood as another feature that hampers the understanding of a model structure. Finally, nonlinear models, like Random Forests, but also neural networks, or nonlinear regression can provide erroneous predictions outside the training data range. This can be especially undesirable if an extreme event is to be predicted.

However, using Random Forests in our study is justified by the following arguments. We look at the model structure by analysing the predictors importance, which, unlike looking at the Random Forests trees structure, gives an easily interpretable insight into the model functioning. Moreover, using as many predictors as in our study design gives additional insight by analysing importance of the water flux sources (rainfall or snowmelt) and their temporal dependencies. Finally, the problem with a proper training data range has to be identified during the model development, as it was the case in our study. For an operational model, this risk should be minimized by selecting the longest possible training period, or by setting up the model for forecasting whether discharge will be above or below certain alarm thresholds.

According to our knowledge, this is the first application of Random Forests for high-resolution discharge forecasting. Several studies used Multi-Layer Neural Networks, Support Vector Regression, Self-organizing Maps or other nonlinear models for discharge forecasting [37–41]. However, Random Forests in river flow related application were used for water level prediction [42], identification of important variables for flood prediction [43] and sediments concentration [44].

4.3. Discharge Forecast and Predictors Importance

Discharge forecasts have lower errors for short forecast times than for the long forecast times, with the errors generally stabilizing for forecast time 7–10 h onwards (Figure 6). This effect can be explained by the fact that short forecast times are strongly dependent on the preceding observed discharge due to autocorrelation (Figure 10). The discharge autocorrelation rapidly decreases onwards lag = 1 h until lag = 10 h when the decrease rate becomes lower (Figure 10). The effect of discharge autocorrelation is

also visible in the predictors importance: the discharge based predictors are clearly more important for the forecast times $t = 1\text{--}7$ h than for the $t = 8\text{--}24$ h in snowmelt and rainfall models (Figure 9), and the rainfall predictors gain more importance onwards lag $l = 7$ h for all forecast times t (Figure 8).

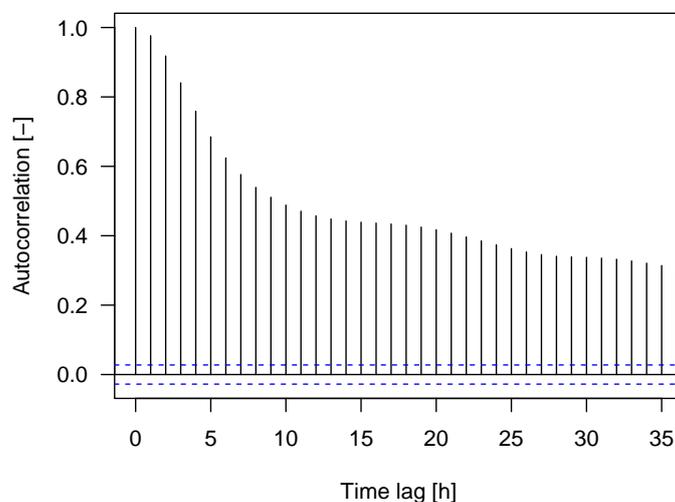


Figure 10. Discharge autocorrelation at the catchment outlet calculated for the calibration and validation period. Horizontal blue lines indicate the autocorrelation significance bounds.

A drawback of the discharge forecasts are some several false positive peaks (Figure 7). As it can be expected, most of these peaks are caused by false positive heavy rainfall events in the WRF forecasts. The magnitude of either false positive rainfall and discharge peaks is small, hence this drawback is not influencing strongly the model performance. In our opinion, these small weather prediction errors are unavoidable; however, they can be minimized by optimizing the WRF setup. This aspect was, however, not in the scope of this study.

Another group of false positive discharge peaks occurring in the snowmelt and rainfall model during an intensive snowmelt event. Unfortunately, we were not able to fully validate the snowmelt flux (e.g., with snow water equivalent data) because only snow depths were available at the meteorological station in the study area; however, we believe that the false positive peaks are a result of overestimated snowmelt in the WRF forecast. Similarly, like for the false positive rainfall, this problem could be diminished by changing the WRF configuration.

Notably, the magnitude of peaks appearing during intensive snowmelt in the snowmelt and rainfall models are lower than the magnitude of peaks appearing by snowmelt followed by rainfall in the rainfall only models. This group of false positive discharge peaks is resulting from lacking snowmelt in the rainfall only models. Regardless, snowmelt being considerably smaller (20% of the total rainfall) than rainfall, its role is clearly identified by the Random Forests models. Snowmelt predictors with relatively higher importance are lagged 17–22 h from the forecast time t . This can be interpreted twofold. First, the high snowmelt predictors' importance at these lags reflects the feedback of snowmelt to create antecedent soil moisture conditions, or initial wetting, that are promoting runoff [20,45,46]. Secondly, it can be a consequence of decreased snowmelt autocorrelation at lags 17 h onwards (Figure 11); however, in this case, one would expect high importance of snowmelt predictors for lags near the prediction time.

The rainfall has increased importance for predictors close to the forecast time t (~0–3 h) and lagged ~12–16 h from the forecast time. The increased importance near the forecast time is due to the fact that Biala River reacts quickly to rainfall, mostly due to extensive sewers and high imperviousness [32,47]. The second period of increased importance matches well with the decreased rainfall autocorrelation onwards 10 h lags (Figure 12). The delayed rainfall is not correlated to the rainfall at the forecast time,

hence it provides new information. It can also have a role in promoting runoff by forming ascendant soil moisture conditions, as we hypothesized above for snowmelt.

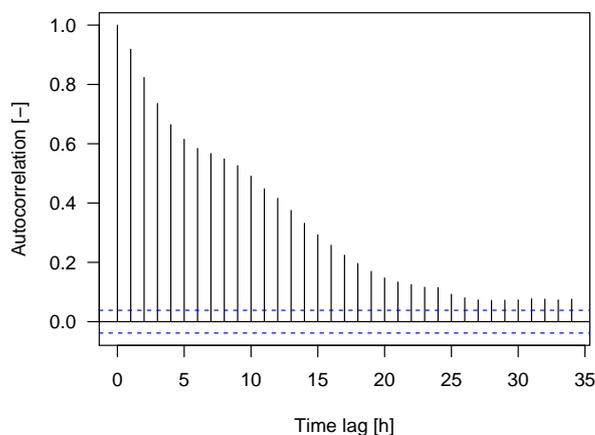


Figure 11. WRF forecasted snowmelt autocorrelation at the catchment outlet calculated for the calibration and validation period. Horizontal blue lines indicate the autocorrelation significance bounds.

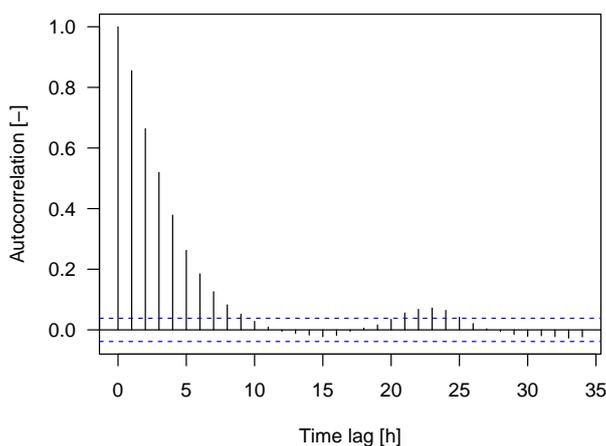


Figure 12. WRF forecasted rainfall autocorrelation at the catchment outlet calculated for the calibration and validation period. Horizontal blue lines indicate the autocorrelation significance bounds.

Despite the snowmelt flux being considerably lower than the rainfall flux, its influence on the forecasted discharge is beneficial. This is clear when comparing importance of discharge based predictors in the snowmelt and rainfall and rainfall only models (Figure 9). As mentioned earlier for the snowmelt and rainfall models, the importance of discharge based predictors decreases onwards forecast times $t = 7$ h. This is not the case for the rainfall only models where the importance of the discharge based predictors is at a similar level at all forecast times. This means that inclusion of snowmelt flux allowed the model to shift the weight from preceding discharge data onto the forcing data, as it is desired.

The effect of the forcing data being more important in the model than the discharge data is expressed in the ME closer to optimum in the snowmelt and rainfall models than in the rainfall only models (Figure 6). The increased bias in the rainfall only models can lead to higher discharge forecast uncertainty, especially during extreme events.

4.4. Snowmelt in Discharge Forecasting

So far, hydrological simulations and forecasts using WRF in snow regime catchments used the meteorological fields (e.g., precipitation, temperature, solar radiation, etc.) to drive snowmelt

simulation in physical [19] and empirical [18] models. Our study shows that driving another model with WRF data is not necessarily crucial to obtaining reliable snowmelt estimates and the WRF–Noah model deriving snowmelt itself can be a valuable estimate.

The majority of the WRF driven hydrological forecasts reported in literature are conducted in sites where snow processes do not occur (e.g., [6,9,48]). However, some studies that lack information of how snowmelt was handled report good hydrological forecasting results in areas where snow processes, including snowmelt, may occur next to or as a mix with rainfall (e.g. [8,42]). Rainfall fields can produce acceptable discharge forecasting in sites where snowmelt is not dominating over rainfall, as shown in our study. However, inclusion of snowmelt in such case studies should not be neglected because it allows the physical processes' representation to be more complete, e.g., in taking into account the antecedent soil moisture conditions [20], rain-on-snow [49] or the runoff [46].

Inclusion of climatic variables into machine-learning models improves the predicted discharge coefficient of determination in reference to models that do not use climatic variables [50]. Moreover, it is known that snowmelt should be included into flood frequency analysis if snow accumulation is substantial [51]. Our results are in agreement with these and point out that, even though snowmelt accounted only for 20% of total rainfall in this study, its effect on the model structure and forecast errors is clearly positive.

4.5. Outlook and Applicability

We believe that the framework of short-term high-resolution discharge forecast presented herein can be adopted in other research studies or as an operational model. One should, however, take care about proper training of the Random Forests models, i.e., with a representative observation period. Another issue is selecting other, hydrological cycle components (e.g., groundwater recharge/discharge, or evapotranspiration) as predictors. A hydrological forecasting with Random Forests model in configuration presented in this study will not be appropriate if an extensively managed water reservoir is operational in a catchment.

We have intentionally selected a relatively small and homogeneous catchment as a study area because this allowed us to use a lumped model and neglect the distributed aspect of snowmelt, precipitation and runoff formation. This simplification highlighted the general importance of snowmelt, rainfall and discharge predictors. However, in our earlier works, we showed that data source (so also the spatial distribution) of snow related variables can considerably change the hydrological modelling results [52]. Moreover, we showed that the spatially distributed snow cover has strongly variable sensitivity in a catchment area [53]. Hence, differentiating the rainfall and snowmelt predictors spatially would be an interesting aspect that would allow to: (1) depict the most important parts of a catchment for runoff formation and (2) potentially improve the forecasts by moving from lumped to distributed modelling.

5. Conclusions

Our study used a nonlinear machine learning algorithm and numerical weather forecasts of snowmelt and rainfall in order to forecast hourly discharge in an urbanized catchment. We tested two scenarios of the discharge forecast model (with rainfall only predictors and with snowmelt and rainfall predictors), which allowed for highlighting the effect of snowmelt predictors in the forecasts. Both scenarios performed similarly in terms of hydrograph behaviour. However, the errors analysis in consecutive forecast hours revealed that the rainfall only models were more biased and had higher absolute errors than the snowmelt and rainfall models. The scenarios comparison showed also that the snowmelt predictors are of comparable importance as the rainfall predictions even though the snowmelt volume is only 20% of the rainfall volume in our study. Moreover, inclusion of snowmelt prediction changed the pattern of preceding discharge predictors by decreasing their importance with increasing forecast lead time.

We conclude that the effect of including snowmelt data in discharge forecasts for mixed snowmelt and rainfall environment allows for accounting for nonlinearities and feedbacks such as (1) initial wetting by snowmelt related to air temperature and (2) rain-on-snow phenomena. The former indicates that antecedent snowmelt is important for discharge forced by precipitation, due to initial wetting and increasing the effective rainfall. The latter indicates that accounting for snowpack properties, such as distribution and water equivalent, is crucial for snow and rainfall originated runoff estimation.

High importance of predictors related to preceding discharge in the rainfall only models showed that the interpretation of the model structure leads to erroneous conclusion (that these predictors are highly important for all forecast lead times in our case) while the predictions had acceptable errors. We showed that the inclusion of snowmelt predictors changed the model's structure in comparison to the rainfall only models in a way that the predictors' importance pattern is interpretable with reference to physical phenomena (e.g., effect of antecedent soil moisture), even though our modelling approach was machine-learning without implementation of any physical equations.

Our approach of discharge forecasting using the Random Forests algorithm has shown its high applicability. We used Random Forests for high-resolution discharge forecasting; however, this approach is similar to other methods, such as neural networks. A clear advantage of our approach is the functionality to measure the predictors' importance, which allows for optimizing the model structure and highlighting its flaws. Thereby, for future work, we are planning to apply the nonlinear forecast models spatially in order to highlight catchment zones that are the most important for discharge forcing.

Acknowledgments: Discharge time series for the Zawady Gauge at the Biala River was provided by the Institute of Meteorology and Water Management National Research Institute (IMGW-PIB). Snow depth and precipitation data for Białystok station were provided by the NOAA Climate Data Online service. Calculations were carried out at the Academic Computer Centre in Gdańsk. This research was supported by Gdańsk University of Technology statutory activity: DS2017 032351. Authors acknowledge two anonymous Reviewers whose comments led to great improvement of this paper.

Author Contributions: Tomasz Berezowski designed and performed the hydrological simulations, analysed the results and prepared the paper; Andrzej Chybicki set up and performed the WRF simulations and revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Buttle, J.M.; Xu, F. Snowmelt Runoff in Suburban Environments. *Hydrol. Res.* **1988**, *19*, 19–40.
- Surfleet, C.G.; Tullos, D. Variability in effect of climate change on rain-on-snow peak flow events in a temperate climate. *J. Hydrol.* **2013**, *479*, 24–34.
- Wever, N.; Jonas, T.; Fierz, C.; Lehning, M. Model simulations of the modulating effect of the snow cover in a rain-on-snow event. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4657–4669.
- Cohen, J.; Ye, H.; Jones, J. Trends and variability in rain-on-snow events. *Geophys. Res. Lett.* **2015**, *42*, 7115–7122.
- Langhammer, J.; Česák, J. Applicability of a Nu-Support Vector Regression Model for the Completion of Missing Data in Hydrological Time Series. *Water* **2016**, *8*, 560.
- Rogelis, M.C.; Werner, M. Streamflow forecasts from WRF precipitation for flood early warning in mountain tropical areas. *Hydrol. Earth Syst. Sci. Discuss.* **2017**, *2017*, 1–32.
- Li, J.; Chen, Y.; Wang, H.; Qin, J.; Li, J.; Chiao, S. Extending flood forecasting lead time in a large watershed by coupling WRF QPF with a distributed hydrological model. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 1279–1294.
- Tao, J.; Wu, D.; Gourley, J.; Zhang, S.Q.; Crow, W.; Peters-Lidard, C.; Barros, A.P. Operational hydrological forecasting during the IPHEX-IOP campaign—Meet the challenge. *J. Hydrol.* **2016**, *541*, 434–456.
- Bauer-Gottwein, P.; Jensen, I.H.; Guzinski, R.; Bredtøft, G.K.T.; Hansen, S.; Michailovsky, C.I. Operational river discharge forecasting in poorly gauged basins: The Kavango River basin case study. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 1469–1485.

10. De Lima, G.R.T.; Santos, L.B.L.; de Carvalho, T.J.; Carvalho, A.R.; Cortivo, F.D.; Scofield, G.B.; Negri, R.G. An operational dynamical neuro-forecasting model for hydrological disasters. *Model. Earth Syst. Environ.* **2016**, *2*, 1–9.
11. Oleson, K.W.; Lawrence, D.M.; Gordon, B.; Flanner, M.G.; Kluzek, E.; Peter, J.; Levis, S.; Swenson, S.C.; Thornton, E.; Feddema, J.; et al. *Technical Description of Version 4.0 of the Community Land Model (CLM)*; National Center for Atmospheric Research: Boulder, CO, USA, 2010.
12. Lawrence, D.M.; Oleson, K.W.; Flanner, M.G.; Thornton, P.E.; Swenson, S.C.; Lawrence, P.J.; Zeng, X.; Yang, Z.L.; Levis, S.; Sakaguchi, K.; et al. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **2011**, *3*, doi:10.1029/2011MS00045.
13. Benjamin, S.G.; Grell, G.A.; Brown, J.M.; Smirnova, T.G.; Bleck, R. Mesoscale Weather Prediction with the RUC Hybrid Isentropic-Terrain-Following Coordinate Model. *Mon. Weather Rev.* **2004**, *132*, 473–494.
14. Niu, G.Y.; Yang, Z.L.; Mitchell, K.E.; Chen, F.; Ek, M.B.; Barlage, M.; Kumar, A.; Manning, K.; Niyogi, D.; Rosero, E.; et al. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.* **2011**, *116*, doi:10.1029/2010JD015139.
15. Wang, Z.; Zeng, X.; Decker, M. Improving snow processes in the Noah land model. *J. Geophys. Res.* **2010**, *115*, D20108, doi:10.1029/2009JD013761.
16. Jin, J.; Wen, L. Evaluation of snowmelt simulation in the Weather Research and Forecasting model. *J. Geophys. Res. Atmos.* **2012**, *117*, doi:10.1029/2011JD016980.
17. Förster, K.; Meon, G.; Marke, T.; Strasser, U. Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany). *Hydrol. Earth Syst. Sci.* **2014**, *18*, 4703–4720.
18. Wu, X.; Shen, Y.; Wang, N.; Pan, X.; Zhang, W.; He, J.; Wang, G. Coupling the WRF model with a temperature index model based on remote sensing for snowmelt simulations in a river basin in the Altay Mountains, north-west China. *Hydrol. Process.* **2016**, *30*, 3967–3977.
19. Zhao, Q.; Liu, Z.; Ye, B.; Qin, Y.; Wei, Z.; Fang, S. A snowmelt runoff forecasting model coupling WRF and DHSVM. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 1897–1906.
20. Buttle, J.M. Effects of suburbanization upon snowmelt runoff. *Hydrol. Sci. J.* **1990**, *35*, 285–302.
21. Valtanen, M.; Sillanpaa, N.; Setälä, H. Effects of land use intensity on stormwater runoff and its temporal occurrence in cold climates. *Hydrol. Process.* **2013**, *28*, 2639–2650.
22. Valeo, C.; Ho, C. Modelling urban snowmelt runoff. *J. Hydrol.* **2004**, *299*, 237–251.
23. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.
24. Instytut Meteorologii i Gospodarki Wodnej—Państwowy Instytut Badawczy (IMGW-PIB). *1 h Discharge Time Series for Zawady (ID: 153230060) Gauging Station at Biala River (ID: 26168) for the Period 2010–2016*; IMGW-PIB: Warsaw, Poland, 2017.
25. Nakanishi, M.; Niino, H. Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer. *J. Meteorol. Soc. Jpn.* **2009**, *87*, 895–912.
26. Mlawer, E.J.; Taubman, S.J.; Brown, P.D.; Iacono, M.J.; Clough, S.A. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.* **1997**, *102*, 16663–16682.
27. Dudhia, J. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *J. Atmos. Sci.* **1989**, *46*, 3077–3107.
28. National Centers for Environmental Information-National Oceanic and Atmospheric Administration (NOAA-NCEI). *Daily Precipitation, Temperature and Snow Depth Time Series for the Bialystok, PL Station (ID: PLM00012295) 2010–2014*; NOAA-NCEI: Washington, DC, USA, 2017.
29. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2016.
30. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.
31. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift* **2006**, *15*, 259–263.
32. Berezowski, T.; Chormański, J.; Batelaan, O.; Canters, F.; Van de Voorde, T. Impact of remotely sensed land-cover proportions on urban runoff prediction. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *16*, 54–65.

33. Kioutsioukis, I.; de Meij, A.; Jakobs, H.; Katragkou, E.; Vinuesa, J.F.; Kazantzidis, A. High resolution WRF ensemble forecasting for irrigation: Multi-variable evaluation. *Atmos. Res.* **2016**, *167*, 156–174.
34. Yu, W.; Nakakita, E.; Kim, S.; Yamaguchi, K. Impact Assessment of Uncertainty Propagation of Ensemble NWP Rainfall to Flood Forecasting with Catchment Scale. *Adv. Meteorol.* **2016**, *2016*, 1384302.
35. García-Díez, M.; Fernández, J.; Fita, L.; Yagüe, C. Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. *Q. J. R. Meteorol. Soc.* **2012**, *139*, 501–514.
36. Chybicki, A.; Łubniewski, Z.; Kamiński, L.; Bruniecki, K.; Markiewicz, L. Numerical weather prediction-data fusion to GIS systems and potential applications. In *The Future with GIS. Katedra Systemow Geoinformatycznych; Krvatski Informsticki Zbor—GIS Forum: Zagreb, Croatia, 2011*; pp. 56–61.
37. Barbetta, S.; Coccia, G.; Moramarco, T.; Todini, E. Case Study: A Real-Time Flood Forecasting System with Predictive Uncertainty Estimation for the Godavari River, India. *Water* **2016**, *8*, 463.
38. Barge, J.; Sharif, H. An Ensemble Empirical Mode Decomposition, Self-Organizing Map, and Linear Genetic Programming Approach for Forecasting River Streamflow. *Water* **2016**, *8*, 247.
39. Peng, T.; Zhou, J.; Zhang, C.; Fu, W. Streamflow Forecasting Using Empirical Wavelet Transform and Artificial Neural Networks. *Water* **2017**, *9*, 406.
40. Sung, J.; Lee, J.; Chung, I.M.; Heo, J.H. Hourly Water Level Forecasting at Tributary Affected by Main River Condition. *Water* **2017**, *9*, 644.
41. Wang, J.; Shi, P.; Jiang, P.; Hu, J.; Qu, S.; Chen, X.; Chen, Y.; Dai, Y.; Xiao, Z. Application of BP Neural Network Algorithm in Traditional Hydrological Model for Flood Forecasting. *Water* **2017**, *9*, 48.
42. Li, B.; Yang, G.; Wan, R.; Dai, X.; Zhang, Y. Comparison of random forests and other statistical methods for the prediction of lake water level: A case study of the Poyang Lake in China. *Hydrol. Res.* **2016**, *47*, 69–83.
43. Albers, S.J.; Déry, S.J.; Petticrew, E.L. Flooding in the Nechako River Basin of Canada: A random forest modeling approach to flood analysis in a regulated reservoir system. *Can. Water Resour. J. Rev. Can. Ressour. Hydr.* **2015**, *41*, 250–260.
44. Francke, T.; López-Tarazón, J.; Schröder, B. Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrol. Process.* **2008**, *22*, 4892–4904.
45. Buttle, J. Soil moisture and groundwater responses to snowmelt on a drumlin sideslope. *J. Hydrol.* **1989**, *105*, 335–355.
46. Bengtsson, L.; Westerström, G. Urban snowmelt and runoff in northern Sweden. *Hydrol. Sci. J.* **1992**, *37*, 263–275.
47. Tyszewski, S.; Kardel, I. *Studium Hydrograficzne Doliny Rzeki Białej z Wytycznymi Do Zagospodarowania Rekreacyjnowypoczynkowego I Elementami Małej Retencji Oraz Prace Hydrologiczne Niezbędne Do Sporządzenia Dokumentacji Hydrologicznej*; Pro Woda: Warsaw, Poland, 2009. (In Polish)
48. Givati, A.; Lynn, B.; Liu, Y.; Rimmer, A. Using the WRF Model in an Operational Streamflow Forecast System for the Jordan River. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 285–299.
49. Harr, R. Some characteristics and consequences of snowmelt during rainfall in western Oregon. *J. Hydrol.* **1981**, *53*, 277–304.
50. Chu, H.; Wei, J.; Li, J.; Qiao, Z.; Cao, J. Improved Medium- and Long-Term Runoff Forecasting Using a Multimodel Approach in the Yellow River Headwaters Region Based on Large-Scale and Local-Scale Climate Information. *Water* **2017**, *9*, 608.
51. Fassnacht, S.R.; Records, R.M. Large snowmelt versus rainfall events in the mountains. *J. Geophys. Res. Atmos.* **2015**, *120*, 2375–2381.
52. Berezowski, T.; Chormański, J.; Batelaan, O. Skill of remote sensing snow products for distributed runoff prediction. *J. Hydrol.* **2015**, *524*, 718–732.
53. Berezowski, T.; Nossent, J.; Chormanski, J.; Batelaan, O. Spatial sensitivity analysis of snow cover data in a distributed rainfall-runoff model. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 1887–1904.

