Accuracy assessment of heavy rainfall forecasting in the Western Caucasus with ICON-Eu regional atmospheric model for short-term flood forecasting

P A Belyakova\(^1\), A N Shikhov\(^3\), S I Perminov\(^4\) and V M. Moreydo\(^1\)

\(^1\)Water Problems Institute of RAS, 119333, 3 Gubkina street, Moscow, Russia
\(^2\)Pacific Geographical Institute of FEB RAS, 690041, 7 Radio street, Vladivostok, Russia
\(^3\)Perm State University, 15 Bukireva street, Perm, 614990, Russia
\(^4\)SCANEX Group, 108811, 22 km of Kievskoe highway, Moscow, Russia

pobel@mail.ru

Abstract. Heavy rainfalls and rain-induced floods regularly cause substantial damage in the Western Caucasus region. In this study, we assess the accuracy of heavy rainfall forecasting with ICON-Eu regional atmospheric model (developed by the weather service of Germany) and the possibility of its use as input data for rain flood prediction. The main conclusion is that the forecast accuracy is strongly determined by the nature of the rainfall in question (mainly convective or triggered by synoptic-scale processes) and the season of the year. The ICON-Eu model systematically underestimates (by 2-3 times) the precipitation amount in the warm season (April - September) and almost never reproduces local convective heavy rainfall events. Therefore, its forecasts for short-term prediction of summertime rain floods have low efficiency. On the other hand, in the cold season (October - March) the model adequately reproduces heavy precipitation events, with some underestimation of the maximum precipitation amount and overestimation of the coverage area. These forecasts can be used to improve the short-term prediction of flash floods on the rivers of the Black Sea coast of the Caucasus during the period from October to March. In addition, we have performed a detailed analysis of the heavy rainfall and a related flood event which occurred on August 17-18, 2019 by comparing hourly observed precipitation with ICON-Eu and Cosmo-Ru forecasts, as well as applying WRF simulation and rainfall-runoff modelling.

1. Introduction

The Western Caucasus is one of the most flood-impacted areas of Russia and the only one where floods regularly cause a substantial number of deaths [1]. The most devastating floods are formed by extreme precipitation during summer months [2]. Under the conditions of non-stationary climate, the probability of extreme precipitation and associated catastrophic floods may increase [3, 4].

Short-term forecasting of rain floods in the Western Caucasus as well as in other mountainous regions is challenging due to a local nature of heavy precipitation events and low density of the precipitation gauging network [5]. In such conditions, weather radar data (see, for example, [6]), short-term forecasting with the use of mesoscale numerical weather prediction (NWP) models [7] or even geostationary meteorological satellite images [8] may be used as sources of initial data to drive runoff formation models. Modern flood forecasting systems use ensemble information from radar-based
prediction and NWP models to improve the forecast accuracy [9]. However, there are some difficulties in obtaining access to weather radar data, which is a specific problem for the Western Caucasus and other Russian regions.

Recently, due to increasing performance of the computing systems, the horizontal grid size of global NWP models has been reduced to 9-13 km. Such grid size, as well as the use of non-hydrostatic approximation allow one to partially resolve mesoscale processes (including deep convection) and use the global NWP model outputs for their forecasting along with mesoscale models [10]. Among the publicly available sources of operational precipitation forecasts, the ICON-Eu model is of greatest interest due to its high spatial resolution (6.5 km). The aim of this study is to assess the accuracy of short-term precipitation forecast according to the ICON-Eu regional NWP model and the possibility of its use as input precipitation data for an event-based rainfall-runoff model KW-GIUh. In addition, we compared the precipitation forecast of the ICON-Eu model with the same of mesoscale NWP models WRF [11] and Cosmo-Ru 7 km [12] for the most important heavy rainfall event that occurred during the study period (17-18 August 2019).

2. Data and Methods

2.1. Icon-Eu NWP model data

The ICON-Eu NWP model is a regional version of the global NWP model ICON operated by the Weather Service of Germany (DWD). The model is launched in operational mode for a domain covering most of the territory of Europe (up to 45° E) from July 2015 [13]. The Western Caucasus region is completely covered by the ICON-EU model data. As compared with the output of other publicly-available NWP models, ICON-Eu data have the highest spatial resolution which is crucial for the prediction of convective precipitation.

We downloaded 27-h operational forecasts of the ICON-Eu model from the FTP server of the DWD (Table 1). The forecast data were obtained twice a day (from 00 h and 12 h UTC). The data have been downloaded for the period from December 2018 to the present. We assessed the forecast accuracy for the period from January 2019 to July 2020. We used Yandex Compute Cloud for the data processing (downloading, cutting and conversion from GRIB2 to GeoTiff format, and resampling to 5000 m cell size with the use of a bilinear interpolation function).

| NWP Model | Developers | GRID resolution, km | Number of vertical levels | Output data grid resolution | Output data time step | Initial conditions for start the model |
|-----------|------------|---------------------|--------------------------|---------------------------|----------------------|--------------------------------------|
| ICON-Eu   | Deutscher Wetterdienst (DWD), Germany (NCEP) and Penn State University, U.S. | 6.5 | 60 | 6.5 km | 1 h | - |
| WRF-ARW   | International Consortium for Small-Scale Modelling | 3.0 | 64 | 3 km | 1 h | ERA-5 reanalysis |
| Cosmo-RU  |                                    | 7 | ? | 7 km | 3 h | ICON-Global NWP model |

2.2. Analysis of heavy rainfall events

We compiled the data on heavy rainfall events (≥ 30 mm/12 h) from Roshydromet 54 weather stations located in the Krasnodar Territory, Adygaya Republic and Stavropol region (within Kuban’ river basin). We analyzed the observed heavy rainfall events for the same period as the ICON-EU forecast data from January 2019 to July 2020. In total, 189 heavy rainfall events have been reported by the weather stations on 66 different days. Heavy rainfall events occurred in all months of the year except April, with the largest number of events in August 2019 (25 events). In July 2019, 11 days with heavy
rainfall have been reported. Among 189 heavy rainfall reports, the criteria of hazardous weather events accepted by Roshydromet (≥ 50 mm/12 h) were exceeded in 61 cases, and twice the precipitation amount exceeded 100 mm/12 h. However, we may underestimate the number of cases with precipitation amount ≥ 50 mm/12 h, since we used only standard 12-h observations of the weather stations (carried out at 03.00 and 15.00 UTC). Thus, heavy rainfall events that did not fall into one observation period are not considered in our data. Also, heavy rainfall could occur between the rare rain gauges and, thus, it could not be taken into analysis. The highest number of heavy rainfall events have been reported on Aug 17, 2019 (11 reports, including 7 ones with precipitation amount ≥ 50 mm/12 h) and on February 4, 2020 (13 reports, including 9 ones with precipitation amount ≥ 50 mm/12 h).

In this study, we considered only the days when the weather stations reported precipitation amount ≥ 30 mm/12 h and did not use any object-based approaches to compare the simulated and observed precipitation amount [14]. First, this is due to the lack of weather radar data. In addition, this is not acceptable for the ICON-Eu model, since it simulates convection with parametrization (as other global NWP models) and does not reproduce mesoscale convective systems explicitly.

To assess the forecast skills of the ICON-Eu model, we used two metrics, such as the Critical Success Index (CSI) and the Extreme Dependency Score (EDI). The SCI index was recommended for accuracy assessment of heavy rainfall precipitation forecasts in [15], and the EDI is widely used to assess the forecast of any low-frequency weather phenomena [16, 17]. It is important to note that the EDI is the simplest metric of the accuracy of forecasts of extreme events, which is independent of the frequency of events. The CSI and EDI metrics were calculated based on the contingency table of predicted and observed heavy rainfall events. The SCI and EDI were calculated as follows:

\[
SCI = \frac{TP}{TP + FP + FN}
\]

\[
EDI = \frac{\log F - \log H}{\log F + \log H}
\]

To calculate the EDI values, it is necessary to preliminarily calculate the TP rate (H) and the F rate (F):

\[
H = \frac{TP}{TP + FN}
\]

\[
F = \frac{FP}{FP + TN}
\]

where TP is the number of true positive forecasts, FN is the number of false positive forecasts, FP is the number of false alarms, and TN is the number of true negative forecasts.

We considered as TP all cases when both observed and predicted precipitation amount at the same weather station exceeded 30 mm. We also added several cases when the predicted precipitation amount was < 30 mm, but it differed from the observed one by < 1.5 times.

Among 66 days when heavy rainfall occurred, at least in 2 cases it caused severe rain floods (28 June 2019 on the Kudepsta River, 17-18 August 2019 on the Khosta River). But only in the last case the precipitation amount and water discharge were measured inside the river basin. For these two days hourly precipitation data were obtained from 15 automated rain gauges of the state meteorological network in Big Sochi region. We compared the accuracy of ICON-Eu forecasts with two mesoscale atmospheric models, WRF and Cosmo-Ru. The main characteristics of mesoscale NWP model data are shown in Table 1. The WRF model with an ARW dynamic core was launched by us on our server, with the use of ERA-5 data as initial and boundary conditions. The model started from 12.00 UTC August 16, 2019, with a 3-km horizontal grid step. The simulation has been performed in one domain (700×700 km grid size). The forecast data of the Cosmo-Ru model were provided by the Hydrometeorological Centre of Russia.

2.3. Hydrological data and modelling
We examined the hydrological situation on the rivers of Big Sochi region on 17-18 August 2019 using 10-minutes water level data obtained from the Automated Flood Monitoring System of the Krasnodar Territory EMERCIT (http://www.emercit.com). According to the data measured at 46 gauges, floods
were reported on all rivers along the Black Sea coast from the Ashe River in the north-west to the Psou River on the state border. On small rivers flowing into the sea, floods mainly occurred on 17 August, the rise of the water level ranged mostly from 30 to 136 cm during 1-2 hours. At larger rivers with more elevated catchments (the Shakh, Sochi, Mzymta, and Psou rivers) high floods were reported on the next day, on 18 August 2019. The highest floods were formed on the Khosta and Kherota Rivers (220 and 198 cm of the level rise with a maximum intensity of 40 cm/10 min and 1.5 m/h), but the level of the hazardous event was exceeded only on the Khosta River.

Therefore, we collected the water discharge of the Khosta River at the Khosta hydrological gauging station (basin area of 98.5 km²) on 17-18 of August 2019 and related precipitation from the Russian Hydrometeorological Service (https://gmvo.sknivh.ru/).

An event-based rainfall-runoff KW-GIUH model was applied to simulate the flood on the Khosta River and to demonstrate the model possibilities for hydrological forecasting. A Kinematic-Wave-based Geomorphological Instantaneous Unit Hydrograph model applies geomorphic stream-order information of the catchment and kinematic-wave theory to analytically determine the travel times for overland and channel flows in a stream-ordering sub-basin system [18, 19, 20]. Here we applied a 2-layer KW-GIUH version considering both the surface- and subsurface-flow processes [19]. For the subsurface flow, Darcy’s law was adopted to estimate the runoff travel time. In applying the instantaneous unit hydrographs for hydrograph simulation, the model deals with temporally non-uniform rainfall through convolution integration of the instantaneous unit hydrographs applied to the rainfall excess of varying intensities with time.

The geomorphic characteristics, the lengths and slopes of sub-catchments and channels of each order were estimated by using the DEM HydroSHEDS (https://www.hydrosheds.org/) and ArcGIS tools. The channel width at the outlet was estimated using satellite images. The overland-flow roughness coefficient, the channel roughness coefficient, and hydraulic conductivity were calibrated against the most severe recent flooding events on 25 June 2015 and 24-25 October 2018.

3. Results and Discussion

3.1. Accuracy assessment of heavy rainfall forecast with ICON-Eu NWP model

To assess the accuracy of heavy precipitation forecasts according to the ICON-Eu model, we compared them with observed precipitation data and calculated the SCI and EDI metrics (Table 2).

| Season            | Model start time | TP | FN | FP | TN | SCI | EDI |
|-------------------|------------------|----|----|----|----|-----|-----|
| All study period  | 00.00 UTC        | 66 | 119| 51 | 3572| 0.28| 0.61|
|                   | 12.00 UTC        | 74 | 112| 71 | 3577| 0.29| 0.62|
| April – September | 00.00 UTC        | 16 | 93 | 22 | 2569| 0.12| 0.43|
| (109 events)      | 12.00 UTC        | 21 | 88 | 11 | 2580| 0.18| 0.53|
| October – March   | 00.00 UTC        | 53 | 26 | 59 | 1083| 0.38| 0.76|
| (80 events)       | 12.00 UTC        | 57 | 24 | 65 | 1096| 0.39| 0.78|

The main conclusion regarding the accuracy of the ICON-Eu model forecasts is that it strongly varies depending on the season of the year due to the different genesis of heavy precipitation. Indeed, in the warm season heavy rainfall events have mainly convective nature, most of them are local and reported simultaneously at 1-2 weather stations. In the cold season heavy precipitation events are mainly associated with synoptic-scale processes, which are more reliably reproduced by the model (although the contribution of convection to their formation is also substantial).

The number of heavy rainfall reports in the warm and cold seasons was 109 and 80, respectively, however, the number of days with heavy rainfall events differed more substantially (44 and 22 days, respectively). The SCI and EDI values indicate unsatisfactory accuracy of the forecasts in the warm season and a substantially higher accuracy in the cold season (Table 3, Figure 1a). The TP rate in the cold season (0.69) is more than four times higher than the same in the warm season (0.17). Thus, more than 80% of the heavy rainfall events in the warm season were missed by the model. The number of
FN in the warm season was several times higher than the number of FP, which also indicates systematic underestimation of the precipitation amount.

The model poorly reproduced local heavy rainfall events (reported only at 1-2 weather stations). In such cases, the forecasted precipitation amount was usually several times lower than the observed one (it usually did not exceed 10 mm). The TP rate for the local events is less than 0.2, but it increases to 0.63 for the events reported at five or more stations simultaneously (Figure 1b). In addition, heavy rainfall events covering a large area in the warm season are reproduced with less accuracy than in the cold season. Thus, the ICON-Eu underestimated the precipitation amount by 2-4 times for a heavy rainfall event on August 17, 2019 (Figure 2a). When comparing this forecast with that obtained according to the COSMO-Ru (Figure 2b) and WRF (Figure 2c) NWP model, it is obvious that the ICON-Eu model underestimated the precipitation amount more substantially than the mesoscale NWP models.

![Figure 1. Monthly distribution of TP rate of heavy rainfall forecast (a) and the dependence of TP rate on the number of simultaneously reported heavy rainfall events (b).](image)

It is also interesting that in the warm season the forecasts obtained from 00 h UTC have a slightly higher accuracy than those obtained from 12.00 UTC. It may be related to the daily distribution of convective heavy rainfall events, which has a maximum in the afternoon. Thus, the forecast lead time for daytime precipitation is 15 h if the model started from 00 h UTC.

In the cold season, the model missed only about a third of all heavy rainfall events. The number of FP more than 2 times exceeded the number of FN. Thus, the predicted number of heavy rainfall events in the cold season is even more than the observed one. However, we cannot conclude that the model systematically overestimates the precipitation amount in the cold season. The highest precipitation amount observed in the cold season (70-80 mm/12 h) was slightly underestimated by the model. More precisely, the model somewhat overestimates the area covered by heavy precipitation (that is the main cause of a large number of false alarms), but underestimates their maximum intensity. A similar effect was previously found for the WRF model with a Kain-Fritsch convection scheme using Perm region as an example [21].

3.2. Rainfall intensity forecasting
Forecasting of rainfall intensity (e.g. hourly precipitation amount) is crucial for rain flood prediction. We obtained data on hourly precipitation amount for August 17-18, 2019 observed at 15 weather stations and compared them with the simulated ones according to the ICON-Eu, COSMO-Ru, and WRF NWP models. Unfortunately, the data of the ICON-Eu and COSMO-Ru were obtained with a 3-h time step, and we had to recalculate the precipitation amount with this time step. However, this does not affect the conclusion that all three NWP models underestimate precipitation intensity by several times (Figure 4). The differences between the observed and simulated precipitation with a 3-h time step is higher than for the accumulated precipitation over 12 hours.

Only the WRF model reproduced the rainfall intensity of up to 100 mm/h at the hydrological gauge Khosta (Figure 4d). It is related to the fact that the WRF model simulates convection explicitly...
(without parametrization), unlike the COSMO-Ru and ICON models. However, the observed and simulated maxima of the rainfall intensity do not coincide in time and location.

Figure 2. Accumulated precipitation amount between 03.00 and 15.00 UTC, August 17, 2019 according to ICON-Eu (a), COSMO-Ru (b), and WRF (c) NWP models. Precipitation observed at weather stations is also indicated.
Figure 3. Accumulated precipitation amount between 15.00 UTC February 3, 2020 and 03.00 UTC February 4, 2020 according to the ICON-Eu NWP models. Precipitation observed at weather stations is also indicated.

Figure 4. Observed and simulated 3-h rainfall intensity at the weather stations Sochi (37099, a), Solokh-Aul (37092, b), Gornaya Karusel 1000 (37102, c), and the hydrological gauge Khosta (82046, d) on 17-18 August 2019.
3.3. Flood simulation

The Khosta watershed has a fourth-order stream network, and its area is 98.5 km$^2$. The mean elevation of the study basin is 420 m, and the maximum elevation is 1116 m. The Khosta watershed is underlaid by limestone and marl, which leads to a wide development of karst processes and karst landforms (caves, etc.) in the watershed. Nevertheless, the runoff formation processes, especially during high floods, can be satisfactorily described by the KW-GIUH model (Figure 5). The values of the parameters were taken as 0.20 for the overland-flow roughness and 0.10 for the channel-flow roughness. The parameters differed in the values of the ratio of the surface-flow region to the total watershed area (also known as partial contributing area), as for the flood on 24-25 October it was taken close to 1 (0.9), while for the flood on 17-18 August 2019 it was estimated as 0.6. Though the precipitation amount was close for both events (96 and 112 mm, 56 and 65 mm causing the main flood, respectively). Considering high uncertainty in the input rainfall data (only 1 rain gauge at the river outlet instead of average precipitation in the basin), the main reason could be differences in the soil moisture at the beginning of the flood event and, thus, different infiltration losses. In October 2018 a heavy rainfall causing a high flash flood occurred after precipitation and following flood formation the day before, but in August 2019 a heavy rainfall occurred after 2 weeks without precipitation.

The quality of flood simulation can be described with these measures. The relative error of the peak discharge is in the range from $-3\%$ to $4\%$. The time of the simulated peak discharge was close or equal to the time of the observed peak discharge considering the hourly time step of the estimates. On 17 August 2019 the time of the observed peak discharge was obtained an hour later than the simulated peak, which represents a belated but abrupt reaction of the dry watershed.

![Figure 5. Observed and simulated water discharge of the Khosta River and observed precipitation at hydrological gauge Khosta, a) on 24-25 October 2018 and b) on 17-18 August 2019.](image)

The KW-GIUH model performs satisfactorily when using observed precipitation, but effective precipitation estimates should also be included. Here we did not apply precipitation forecast data from the ICON-Eu or COSMO-Ru models for flood simulation in the Khosta River basin. The forecasted precipitation amount on 17 August 2019 was underestimated by more than 2 times for the study river basin (20 mm by the ICON-Eu, 26 mm by the COSMO-Ru, 65 mm of observed precipitation during 6 hours). Although the heavy rainfall in October 2018 was well forecasted in the Big Sochi region, this allowed one to run the KW-GIUH model for hydrological forecasting on the catchment scale.

Despite major efforts in the detection and short-term forecasting of precipitation with high spatial and temporal resolution, forecasting of flash floods remains challenging [22]. Real-time radar
composites and radar nowcasting could improve flash flood detection, especially in the warm season. The distributed version of the KW-GIUH model is able to adopt such input data [23].

4. Conclusions
In this study, we estimated the accuracy of short-term (27-h) forecasts of heavy rainfall in the Western Caucasus with the regional ICON-Eu NWP model and the possibility of its use as input data for rain flood prediction. The main conclusion is that the ICON-Eu model systematically underestimates (by 2-3 times) precipitation amount in the warm season (April – September) and almost never reproduces local convective heavy rainfall events (reported at 1-2 weather stations). Therefore, its forecasts for short-term prediction of rain flood events have low efficiency. It is also important that the ICON-Eu model underestimates by several times the maximum intensity of precipitation, on the example of a heavy rainfall event that occurred on August 17, 2019.

In the cold season (October – March), the ICON-Eu model adequately reproduces most heavy precipitation events, with some underestimation of the maximum precipitation amount and overestimation of the coverage area (which is expected, taking into account the use of convective parametrization in the model). These forecasts can be used to improve short-term forecasting of rain floods on the rivers of the Black Sea coast of the Caucasus during the period from October to March.

In the warm season (April – September), unsatisfactory accuracy of short-term forecasting of local convective rainfall does not allow using the ICON-Eu model data as meteorological forcing for the hydrological model KW-GIUH. The assimilation of real-time radar data could improve predictability of flash floods, since the model is able to satisfactorily simulate floods by using observed precipitation. Further analysis of observed and simulated floods is needed for more sustainable estimation of the KW-GIUH model parameters.

Acknowledgments
This study was supported by RFBR and the Krasnodar Territory Government under research project no. 19-45-233007.

References
[1] Alekseevskii N I, Magritskii D V, Koltermann P K, Toropov P A, Shkolnyi D I and Belyakova P A 2016 Water Resources 43(1) 1
[2] Belyakova P A and Gartsman B I 2018 Water Resour 45 50
[3] Semenov V A 2011 Russ. Meteorol. Hydrol. 36(2) 124
[4] Meredith E P, Semenov V A, Marau D, Park W and Chernokulsky A V 2015 Nat. Geosci. 8 615
[5] Lure P M and Panov V D 2011 Russ. Meteorol. Hydrol. 36(4) 273
[6] Carpenter T M, Sperfslage J A, Georgakakos K P, Sweeney T and Fread D L 1999 J. Hydrol. 224(1-2) 21
[7] Vincendon B, Ducrocq V, Nuiissier O, Vié B 2011 Nat. Hazards Earth Syst. Sci. 11(5) 1529
[8] Wardah T, Abu Bakar S H, Bardossy A and Maznorizan M 2008 J. Hydrol. 356(3-4) 283
[9] Yu W, Nakakita E, Kim S and Yamaguchi K. 2015 J. Hydrol. 531 494
[10] Rivin G S et al 2019 Russ. Meteorol. Hydrol. 44(11) 729
[11] Powers J G et al 2017 Bull. Amer. Met. Soc. 98 1717
[12] Rivin G S et al 2015 Russ. Meteorol. Hydrol. 40(6) 400
[13] ICON Model Tutorial (April 2019) URL: https://code.mpmet.mpg.de/attachments/download/19568/ICON_tutorial_2019.pdf (assessed 16.08.2020) [14] Davis C A, Brown B and Bullock R 2006 Mon. Wea. Rev. 134 1772
[15] Kiselnikova V Z 2013 Russ. Meteorol. Hydrol. 38(4) 217
[16] Murav’ev A V, Kiktev D V, Bundel’ A Yu, Dmitrieva T G and Smirnov A V 2015 Russ. Meteorol. Hydrol. 40(9) 584
[17] Kiktev D et al 2017 Bull. Amer. Meteorol. Soc. 98(9) 1908
[18] Lee K T and Yen B C 1997 J. Hydraul. Eng. 123 73
[19] Lee K T and Chang C H 2005 J. Hydrol. 311 91
[20] Lee K T, Cheng N K, Gartsman B I and Bugayets A N 2009 Geogr. Nat. Resour. 30(1) 79–85
[21] Kalinin N A, Shikhov A N, Bykov A V and Tarasov A V 2019 Hydromet. studies and forecasting 373(3) 43 (in Russian)
[22] Raynaud D, Thielen J, Salamon P, Burek P, Anquetin S and Alfieri L 2014 Meteorol. Appl. 22 410
[23] Gonchukov L V, Bugaets A N, Gartsman B I and Lee K T 2019 Water Resour 46 S25