TOPIC 2: HYDROLOGICAL DATA MANAGEMENT

EXPECTED CHANGES IN THE 100-YEAR RIVER DISCHARGE IN THE 21ST CENTURY AT THE DANUBE RIVER IN BRATISLAVA

Ladislav Gaál[1], Danica Lešková[2], Eva Uhliarová[2]

[1] MicroStep-MIS s.r.o., Bratislava, **Slovakia**

[2] Slovak Hydrometeorological Institute, Bratislava, **Slovakia**

Corresponding author: Ladislav Gaál, MicroStep-MIS s.r.o., Cavojskeho 1, 841 04 Bratislava, Slovakia, *ladislav.gaal@microstep-mis.com*

ABSTRACT

The paper presents first results of the local case study 'Flood warnings in a changing climate' carried out within the framework of the SWICCA project.

We assess the changes in Q_{100} (i.e., the 100-year return value of river discharge or simply the 100-year flood) that might be expected to occur due to climate change during the 21st century. The target location is the Danube River in Bratislava, Slovakia. We define a simple indicator CCQ_{100} termed as *Climate Change Indicator of the 100-year Flood*, which is the ratio of the 100-year floods estimated on the basis of future and past datasets. For the past, the observed data from the period of 1984–2014 are used. For the future (2015–2100), a number of simulated time series of river discharge from the SWICCA database are processed. They are the outcomes of the combination of 3 hydrological models (HYPE, VIC and Lisflood) and 11 regional climate models.

The case study is an at-site and basically a non-stationary frequency analysis. The data samples are derived on the basis of the block maxima approach (annual maxima), and the flood quantiles are modeled by the Generalized Extreme Value (GEV) distribution. In case there is a significant linear trend in the data samples, the non-stationary approach is adopted with a time-dependent location parameter of the GEV distribution. Flood quantiles and their confidence intervals are estimated on the basis of the Differential Evolution Markov Chain approach.

The results indicate that: (a) The overall performance of the datasets are rather balanced: increase (decrease) in Q_{100} appears in 13 (12) cases, while no change ($0.95 \leq CCQ_{100} \leq 1.05$) is found in 7 cases; (b) Increases dominate in the case of hydrological models HYPE and Lisflood; (c) The HYPE model indicates the largest positive changes overall, also with the highest value of CCQ_{100} (= 1.40); and (d) The VIC model only yields decrease in Q_{100}. The most remarkable drop is of about 33% (CCQ_{100} = 0.67).

Overall, the multitude of model outcomes is an excellent basis to get the first sight on the possible range of expected changes in the 100-year flood. On the other hand, at this stage of the analysis one cannot arrive to a clear conclusion concerning the sign of these changes. It is believed that further information will be obtained by the adoption of a frequency analysis on the basis of the peaks-over-threshold methodology. This, hopefully, will reveal more insight into the expected behavior of floods with low probability of occurrence, and the knowledge might be transformed into flood management and adaptation plans of the capital city of Slovakia.

Keywords: frequency analysis, climate change, 100-year flood, Danube, SWICCA

1. INTRODUCTION

Flood frequency analysis is a part of the operational services of the Slovak Hydrometeorological Institute (SHMI). So far, estimation of *T*-year return values of floods has generally been based on local data (at-site approach), using principle of the stationarity of environment, and analyzing annual or seasonal maxima. While in the last couple of years, regional approaches to a flood frequency analysis have been successfully adopted in Slovakia [1] and the Central European region [2], no methods accounting for the non-stationarity of environment were developed and implemented in the practice. Similarly, little efforts have been done with implementing the peaks-over-threshold (POT) methodology (where all independent events exceeding a pre-defined threshold are taken into account; e.g., [3]), which is generally considered as a more rigorous statistical approach when compared to the block maxima method (e.g., [4]).

Therefore, MicroStep-MIS and SHMI co-operate on the local case study ‘Flood warnings in a changing climate’, which aims at developing flood frequency estimation methods with two novel directions. First, the non-stationarity of environment will be accounted for, and secondly, the return levels of river discharges will be estimated using the POT methodology. Nevertheless, before dealing with the POT method, an interim step is adopted where flood quantiles are assessed on the basis of the annual maxima series (AMS) approach – and the result of this procedure are presented in the current study.

The case study ‘Flood warnings in a changing climate’ is being carried out under the framework of the SWICCA project (http://swicca.climate.copernicus.eu/). SWICCA (**S**ervice for **W**ater **I**ndicators in **C**limate **C**hange **A**daptation) is more than a two-year project governed by the Swedish Meteorological and Hydrological Institute, and serves as a proof-of-concept for a Sectorial Information Service on water management to Copernicus Climate Change Services. The goal of the SWICCA project is to offer freely available climatological and hydrological data (climate change indicators) collected from a number of Pan-European climate/hydrological model runs, to facilitate working with climate change adaptation in the water sector and the decision-making process of water managers. The transfer of the information from the global to regional and/or local scales are demonstrated by means of a number of local case studies from different regions across the whole Europe.

2. METHODS

The presented study aims at assessing the changes in the 100-year return value of river discharge (or simply the 100-year flood, Q_{100}) that might be expected to occur due to the climate change during the 21st century compared to what was observed during the past couple of decades. We define a simple indicator CCQ_{100} termed as *Climate Change Indicator of the 100-year Flood*:

$$CCQ_{100} = \frac{Q_{100,future}}{Q_{100,past}} \quad (1)$$

which is the ratio of the 100-year floods estimated on the basis of the future and the past datasets. For the past, the observed data are used, as usual. For the future, a number of simulated time series of river discharge from the SWICCA database are processed.

Since the outputs of regional climate models are generally affected by bias (errors due to conceptualization, discretization and spatial averaging), bias correction has to be applied to make the model outputs more similar to the reality. For this reason, the so called linear scaling methodology of bias correction [5] is adopted that makes use of the monthly statistics (averages and standard deviations) derived from the common period where both the observed and the modelled discharge data are available.

The basis of the case study is an at-site and non-stationary frequency analysis. The data samples are derived according to the block maxima approach (namely, the annual maxima are identified for each year), and the flood quantiles are statistically modeled by the Generalized Extreme Value (GEV) distribution [6]. The rigorousness of the AMS/GEV approach is justified by the extreme value theory (e.g., [7]).

The stationarity of the time series is analyzed by means of the Mann-Kendall test for the presence of monotonic trends [8] at the significance level of $\alpha = 0.05$. In the case there is a significant linear trend in the data samples, the non-stationary approach to a frequency estimation is adopted with a time-dependent location parameter of the GEV distribution.

Flood quantiles and their confidence intervals are estimated on the basis of the Differential Evolution Markov Chain (DEMC) approach [9]. DEMC is an enhanced alternative to the Markov Chain Monte Carlo (MCMC) approaches where target posterior distributions are sampled through five Markov Chains constructed in parallel; however, in DEMC, the chains are allowed to learn from each other, and this feature ensures simplicity, speed of calculation, and convergence over the conventional MCMC [10].

3. DATA

3.1 Observed data

The presented frequency analysis focuses at a single target site, which is Bratislava, the capital city of Slovakia along the Danube River. The daily discharge data from the Bratislava station (with geographical co-ordinates 48.1397 N, and 17.1082 E) cover the period from January 1st, 1984 till December 31st, 2014, i.e., 31 complete calendar years with no missing values are available for the analysis.

3.2 Simulated data

From the SWICCA database, simulations of daily discharge data were downloaded for the grid box (48.14 N, 17.11 E) corresponding to the location of the Bratislava station, and for the combination of 11 climate (Tab. 1) and 3 hydrological models (Tab. 2).

Since the real observations are from the period 1984–2014, and the control period of the modelled data is 1971–2000, we decided to use the 17 years long common period 1984–2000 as the basis to derive the statistical characteristics for the bias correction. In line with this decision, the period 2015–2100 was declared as the ‘future’.

Tab. 1: Summary of the climate model runs used in the SWICCA database (according to http://swicca.climate.copernicus.eu/wp-content/uploads/2016/10/Metadata_RiverFlow.pdf). GCM = Global Circulation Model, RCM = Regional Climate Model, RCP = Representative Concentration Pathway.

Institute	GCM	RCM	Period	RCP		
				2.6	4.5	8.5
KNMI	EC-EARTH	RACMO22E	1951-2100		✓	✓
SMHI	EC-EARTH	RCA4	1970-2100	✓	✓	✓
	HadGEM2-ES	RCA4	1970-2098		✓	✓
IPSL	CM5A	WRF33	1971-2100		✓	
CSC	MPI-ESM-LR	REMO2009	1951-2100	✓	✓	✓

Tab. 2: Summary of the hydrological models used in the SWICCA database.

Short name	Full name	Information
HYPE	E-HYPE 2.1	http://hypecode.smhi.se/
VIC	VIC-4.2.1.g	http://vic.readthedocs.io/en/master/
Lisflood	Lisflood	[11]

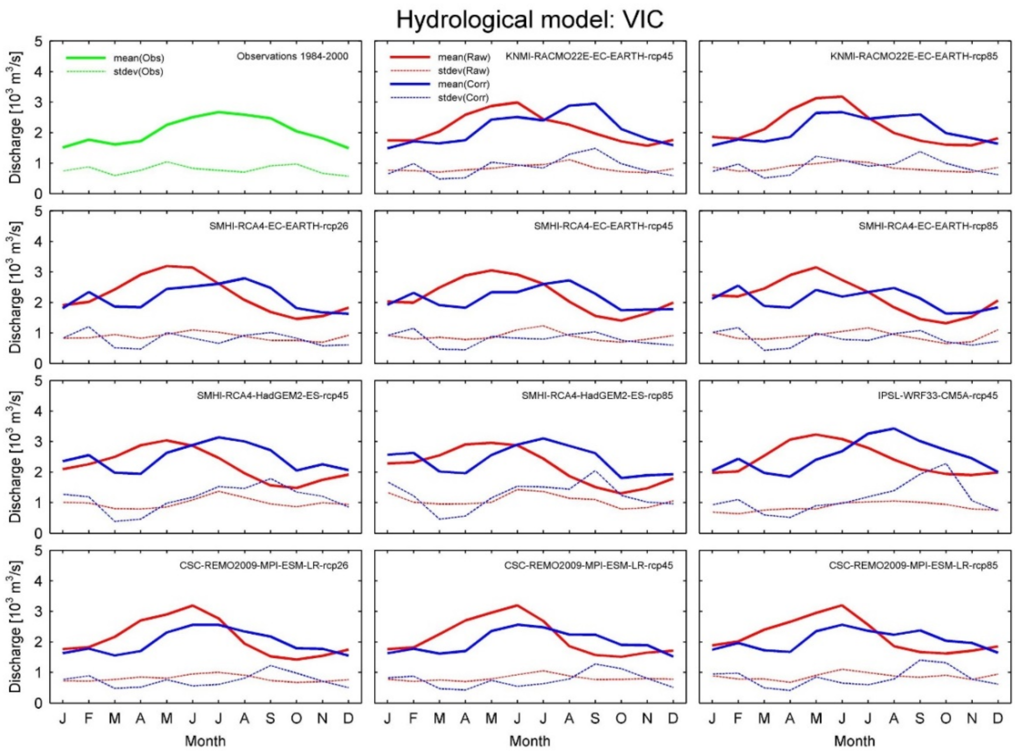

Fig. 1: Monthly characteristics of river discharge for the hydrological model VIC for the future period 2015–2100. The individual plots indicate monthly means (solid, thick lines) and monthly standard deviations (dashed, thin lines) of raw discharge data (red color) and the bias corrected discharge data (blue color). The plot in green color in the top left corner shows the same statistical characteristics for the observed discharge data for the period 1984–2000.

4. RESULTS

4.1 Bias correction

The 33 time series of simulated discharge data were bias corrected on the basis of the linear scaling methodology [5] using the statistical characteristics from the common period 1984–2000. The results of the bias correction for the data series corresponding to the hydrological model VIC are shown in Fig. 1. The comparison of the raw and the bias corrected data shows different patterns that is not straightforward to generalize. Nevertheless, in a number of cases, the bias correction reduced the scale of the discharge values (see the bottom line of the composition in Fig. 1). In other cases, the bias-corrected data show much realistic seasonality, i.e., the annual maxima were pushed from the spring towards the summer months, and at the same time, the annual minima were moved from autumn to the early spring. Similar overall patterns were obtained for the other two hydrological models (not shown here).

4.2 Frequency estimation of the observed data

The frequency estimation using the AMS/GEV approach for the observed discharge data from the period 1984–2014 was carried out. The results of the analysis are presented in Fig. 2. It can be seen that although there is an increasing linear trend in the annual maxima series (Fig. 2, left), it is not significant at level $\alpha = 0.05$. Therefore, the stationary frequency analysis was adopted. The quantile-quantile plot between the empirical quantiles and the theoretical ones corresponding to the GEV distribution function (not shown here) indicates that the GEV distribution is acceptable for modelling the flood quantiles nearly in the entire range of discharges, perhaps with the exception of the largest extremes. The frequency plot (Fig. 2, right) shows the median and the 90% confidence intervals of the return level estimates from a large number of DEMC samples. The higher degree of uncertainty (i.e., wide confidence intervals) is clearly the consequence of the shortness of the analyzed AMS series (31 years).

4.3 Frequency estimation of the simulated data

Similarly as in the case of observed data, the frequency analysis was carried out for the combination of 3 x 11 bias corrected datasets for the future period 2015–2100. Graphical outputs of the frequency analysis for two selected datasets are shown in Fig. 3.

The Mann-Kendall test rejected the null hypothesis about the presence of a linear trend at the significance level $\alpha = 0.05$ in the majority of the cases, i.e., 31 times. There were only two datasets (both related to the Lisflood hydrological model) where a significant linear trend was indicated. In these two cases, the non-stationary approach to a frequency analysis was adopted (one of them is shown in Fig. 3). As in the case of the observed data, the quantile-quantile plots also confirm the applicability of the GEV distribution as a theoretical distribution function for the statistical modelling of the flood quantiles (not shown here). Finally, the frequency plots in all cases (Fig. 3, bottom) clearly show lower degree of uncertainty of the estimated flood quantiles as a result of larger data samples (86 years in most cases).

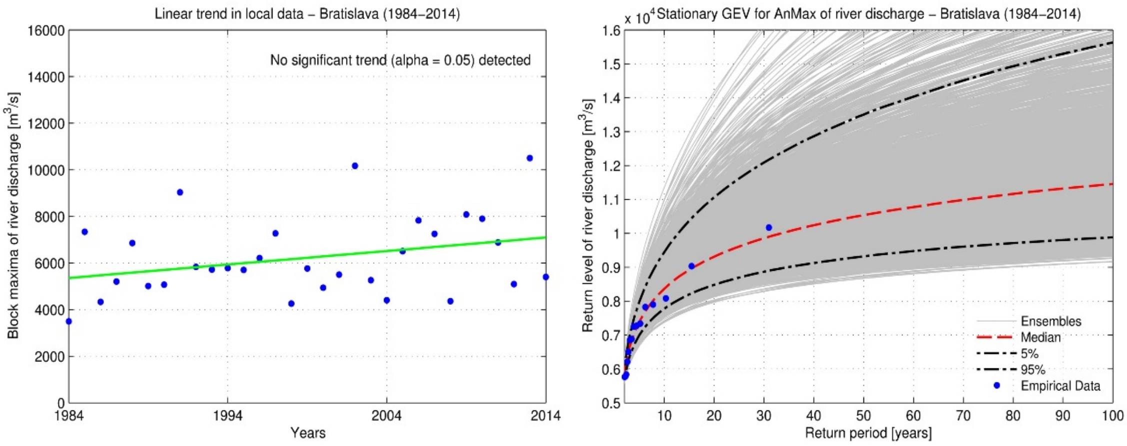

Fig. 2: Annual maxima of river discharge with the trend line (left) and return levels of river discharge on the basis of the AMS/GEV frequency analysis (right) for the observed data of the Bratislava station from the period 1984–2014. In right, the black dash-dotted lines denote the 90% confidence interval for the estimated return levels indicated in red

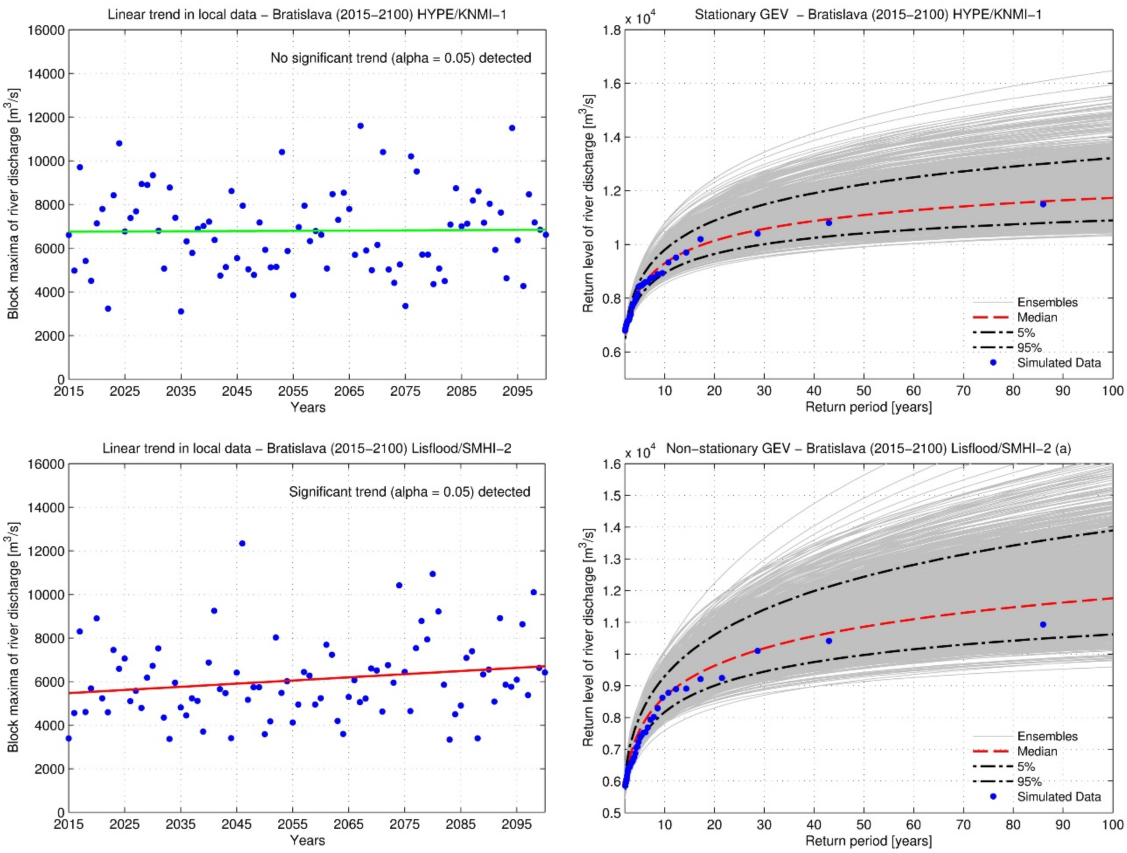

Fig. 3: Annual maxima of river discharge with the trend line (left) and return levels of river discharge on the basis of the AMS/GEV frequency analysis (right) for selected simulated datasets from the period 2015–2100 for Bratislava. Top: the HYPE hydrological model with the KNMI-RACMO22E-EC-EARTH-rcp45 climate model; Bottom: the Lisflood hydrological model with the SMHI-RCA4-EC-EARTH-rcp45 climate model.

4.4 Climate Change Indicator

The results of frequency analysis on the basis of the AMS/GEV approach are summarized in Fig. 4. Displayed are the Q_{100} estimates with their 90% confidence intervals from all 33 simulated SWICCA datasets along with the estimate based on the observed data. Instead of showing the values of the climate change indicator CCQ_{100} themselves, we show the color coded Q_{100} estimates from the future in a relation to the 'real' one, indicated by black circle and the horizontal dashed line in each plot (Fig. 4). We decided to use three qualitative categories (color coding):

- grey color is used for the cases where practically no change in Q_{100} is observed, i.e., the change in absolute value is less than 5% ($0.95 \leq CCQ_{100} \leq 1.05$);
- red color indicates considerable increase in Q_{100} ($CCQ_{100} > 1.05$); and
- green color indicates considerable decrease in Q_{100} ($CCQ_{100} < 0.95$).

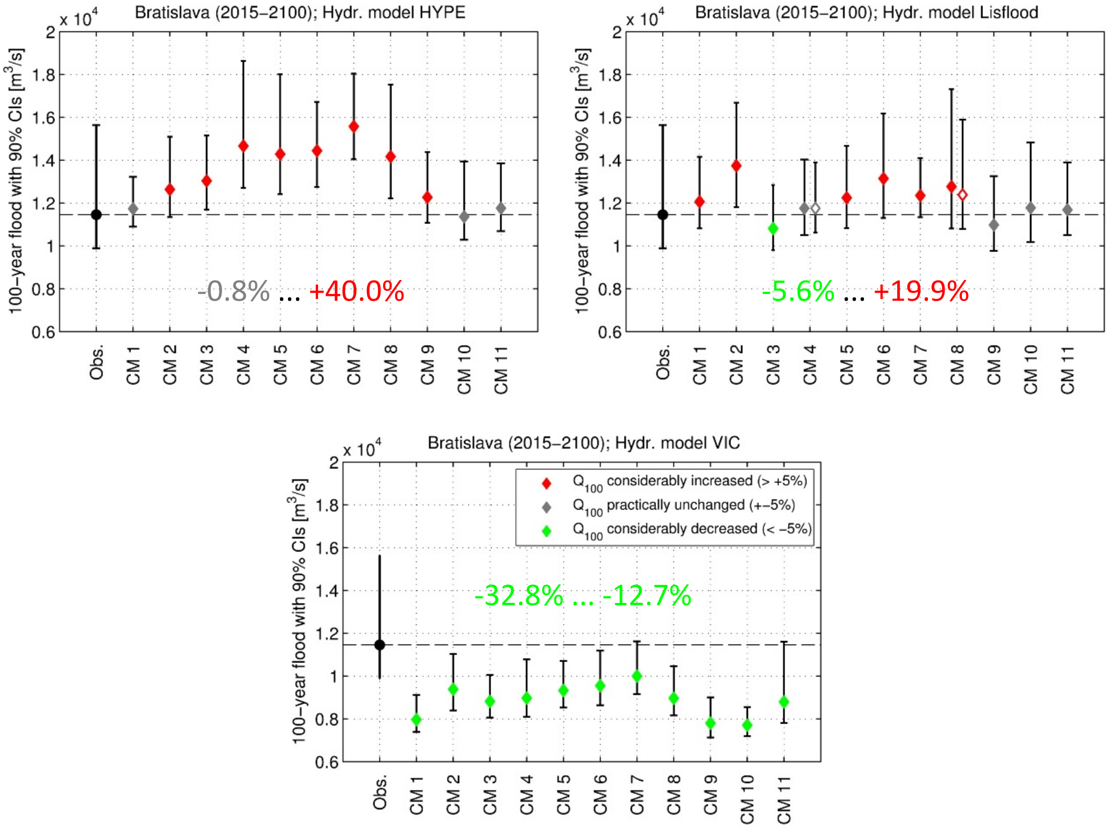

Fig. 4: Estimates of the 100-year flood (colored diamonds) with 90% confidence intervals (whiskers) on the basis of the combination of 3 hydrological models and 11 climate models (abbr. as 'CM 1' to 'CM 11', as they appear in Tab. 1) from the future period 2015–2100. Furthermore, in each plot, the estimate of the 100-year flood on the basis of the observed data from the period 1984–2014 is shown at the very first position (black circle). The horizontal dashed line also corresponds to this estimate. Red (green) color indicate increase (decrease) of a magnitude of >5% (<-5%) in the Q_{100}, while grey color indicates practically no change in Q_{100} (≤±5%). The empty diamonds (two occurrences in the case of the Lisflood model) correspond to the non-stationary approach to a frequency estimation

Fig. 4 reveals interesting features:

- Increases dominate in the case of two hydrological models (HYPE and Lisflood). The difference between these two models lies in the magnitude of the increase: the HYPE model indicates the largest positive changes overall.
- The highest value of CCQ_{100} (= 1.40) appears in the case of the HYPE hydrological model and the SMHI-RCA4-HadGEM2-ES-rcp85 climate model.
- The VIC model only yields decrease in Q_{100}.
- In rough approximation, the patterns of the change in Q_{100} are similar for the hydrological models HYPE and VIC. The largest (smallest) Q_{100} appears both at CM7 - SMHI-RCA4-HadGEM2-ES-rcp8 (CM10 - CSC-REMO2009-MPI-ESM-LR-rcp45), and such an analogy holds for a number of the climate models when comparing HYPE vs. VIC models. From this perspective, the Lisflood hydrologic model shows a fuzzier pattern.
- The overall performance of the datasets are rather balanced: increase (decrease) appears in 13 (12) cases, while no change is indicated in 7 cases (see also Tab. 3).
- Results of the non-stationary approach are also displayed; these are indicated by empty diamonds in the case of Lisflood model. It can be concluded that there are negligible differences between the corresponding results related to stationary vs. non-stationary approach. The variability among the different climate models is much larger than that stemming from the assumption of stationarity or non-stationarity.

Tab. 3 summarizes the most important statistics of the analysis.

Tab. 3: Summary of the results of the frequency analysis based on the approach AMS/GEV. CCQ_{100} stands for the Climate Change Indicator of the 100-year Flood.

Hydrol. model	Climate models	CCQ_{100} > 1.05	CCQ_{100} = 1.05…0.95	CCQ_{100} < 0.95	Largest increase	Largest decrease
HYPE	11	8	3	0	40%	-1%
VIC	11	0	0	11	---	-33%
Lisflood	11	5	4	1	20%	-6%
All	**33**	**13**	**7**	**12**	**40%**	**-33%**

5. DISCUSSION AND CONCLUSIONS

The current paper presents an analysis of the expected changes in estimates of the 100-year flood during the 21st century at the selected target site, Bratislava. The analysis made use of the database of the SWICCA project that consists of a wide variety of climate and hydrological model runs that are available for the upcoming decades from different international projects and databases. In the next paragraphs, we are first going to discuss some particular settings of the analysis that are expected to have influenced our results. Later, a more general evaluation of the benefits and the drawbacks of the concept based on the SWICCA climate indicators will be given.

We are aware of the fact that one of the limiting factors of the analysis is the shortness of the observed data series (1984–2014). This is directly represented by considerably wide confidence intervals of the return level estimates. Furthermore, the shortness of the observed data series influenced the definition of the common period to derive the statistical characteristics of the observed and the modelled data for the bias correction. We had to restrict ourselves to a period of a length of 17 years (1984–2000).

The selected method of bias correction might have also influenced the outcomes. Since we are focusing on extremes, it may be more rigorous to apply a more sophisticated bias correction method, such as one based on the similarity of the empirical distribution functions (e.g., the method of 'distribution mapping' in [5]).

Generally, it is positive that one does not have to run complex hydrological models locally to get future hydrological data and indices; instead, SWICCA offers easy access to these data. SWICCA allows for (especially in the case of the current case study) getting 11x3 time series of river discharge, which are used to estimate the design 100-year discharge for the 21st century for the target site. The analysis results in 33 estimates of Q_{100} in a relatively wide range, which is beneficial since they correspond to a wide diversity of emission scenarios, global and regional climate models and hydrological models. On the other hand, the qualitative results (i.e., whether the Q_{100} is expected to increase/decrease) highly depend on the particular hydrological model. In other words, the three hydrological models translate the same set of 11 regional climate model inputs into considerable different hydrological outputs. This fact emphasizes the uncertainties hidden in the hydrological models, so one has rather to avoid model-based interpretations of the outcomes.

The working hypothesis is that the improved statistical methods of frequency analysis with the combination of the SWICCA climate impact indicator will reveal more insight into the expected behavior of floods with low probability of occurrence, and this knowledge might be transformed into flood management and adaptation plans.

The boundary banks of the Danube River in Bratislava are designed on the basis of the estimate of Q_{100}, and are constructed with a sufficient reserve (reliability) to resist against even larger floods. The information on Q_{100} on the basis of 33 scenarios together with the knowledge on Q_{100} from the past decades are useful at least from two aspects: (i) from the qualitative point of view, i.e., one can see what percentage of the scenarios yield considerable increase/decrease/no change, and (ii) from the quantitative point of view, i.e., one can assess the recent status of the flood prevention system, both in the light of the worst and the best scenarios. The largest values of CCQ_{100} may directly and indirectly indicate the amount of necessary investments (financial, material, logistical, political etc.) into the flood prevention system. On the other hand, even the best scenarios (cases with $CCQ_{100} < 0.95$) may convey valuable information. In this case the buildings that have been constructed on the basis of 'old' estimates of Q_{100} would not need to be rebuilt, they may be declared as flood safe, eventually as protected even against the 1000-year flood. Furthermore, some of the new constructions (dams, bridges) will have lower costs of realization and running expenses.

Overall, the multitude of the model outcomes is an excellent basis to get the first sight on the possible range of the expected changes in the 100-year flood. On the other hand, at this stage of the analysis one cannot arrive to a clear conclusion concerning the sign of these changes. Generally, it is expected that adoption of the novel frequency estimation approach (peaks-over-threshold method) and its comparison with the current AMS/GEV approach will shed more light on the unresolved problems. It is anticipated that POT method will show more beneficial statistical behavior (i.e., narrower confidence intervals of the return levels); nevertheless, this analysis is still to be carried out.

ACKNOWLEDGEMENTS

The support of the SWICCA project is gratefully acknowledged.

REFERENCES

[1] Gaál, L., Kyselý, J., Szolgay, J.: Region-of-influence approach to a frequency analysis of heavy precipitation in Slovakia. *Hydrology and Earth System Sciences*, **12** (3), 825–839, 2008. doi:10.5194/hess-12-825-2008.

[2] Gaál, L., Kyselý, J.: Comparison of region-of-influence methods for estimating high quantiles of precipitation in a dense dataset in the Czech Republic. *Hydrology and Earth System Sciences*, **13** (11), 2203–2219, 2009. doi:10.5194/hess-13-2203-2009.

[3] Bačová-Mitková, V., Onderka, M.: Analysis of extreme hydrological events on the Danube using the peak over threshold method. *Journal of Hydrology and Hydromechanics*, **58** (2), 88–101, 2010. doi:10.2478/v10098-010-0009-x.

[4] Madsen, H., Rasmussen, P.F., Rosbjerg, D.: Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1: At-site modeling. *Water Resources Research*, **33** (4), 747–757, 1997. doi:10.1029/96WR03848.

[5] Teutschbein, C., Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, **456–457**, 12–29, 2012. doi:10.1016/j.jhydrol.2012.05.052.

[6] Coles, S.: *An introduction to statistical modeling of extreme values*. London : Springer, 2001.

[7] Katz, R.W., Parlange, M.B., Naveau, P.: Statistics of extremes in hydrology. *Advances in Water Resources*, **25**, 1287–1304, 2002. doi:10.1016/S0309-1708(02)00056-8.

[8] Wilks, D.S.: *Statistical methods in the atmospheric sciences*. Oxford : Academic Press, 2011. 3rd ed.

[9] Cheng, L., AghaKouchak, A., Gilleland, E., Katz, R.W.: Non-stationary extreme value analysis in a changing climate. *Climatic Change*, **127**, 353–369, 2014. doi:10.1007/s10584-014-1254-5.

[10] Cheng, L., AghaKouchak, A.: Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific Reports*, **4**, 7093, 2014. doi:10.1038/srep07093.

[11] Burek, P.A., van der Knijff, J., de Rio, A.: LISFLOOD – Distributed Water Balance and Flood Simulation Model - Revised User Manual 2013. Joint Research Centre, Report EUR 26162 EN. Luxembourg : Publication Office of the European Commission, 2013.