The probabilistic hydrological model MARCS^{HYDRO} (MARkov Chain System): structure and the core version 0.2

Elena Shevnina[1], Andrey Silaev[2]

[1]Finnish Meteorological Institute, Helsinki, FI-00560, Finland
[2]National Research University Higher School of Economics, Nizhny Novgorod, 603155, Russia.

Correspondence to: Elena Shevnina (elena.shevnina@fmi.fi)

Abstract. A question of environmental risks of social and economic infrastructure has become apparent recently due to an increase in the number of extreme weather events. Extreme runoff events include floods and droughts. In water engineering extreme runoff is described in terms of probability, and uses methods of frequency analysis to evaluate an exceedance probability curve (EPC) of runoff. It is assumed that historical observations of runoff are representative for the future; however trends in observed time series doubt this assumption. The paper describes a probabilistic hydrological model Markov Chain System (MARCS^{HYDRO}) to be applied to predict future runoff extremes. The MARCS^{HYDRO} model simulates statistical estimators of a multi-year runoff to perform future projections in probabilistic form. Projected statistics of meteorological variables available in climate scenarios force the model. This study introduces a new model's core version, and provides its user guide together with an example of the model set up for a single case study. In this case study, the model simulates projected exceedance probability curves of annual runoff under three climate scenarios. The scope of applicability and limitations of the model's core version 0.2 are discussed.

Introduction

Streamflow runoff serves water resources for humans, food production and energy generation, while risks of water-sensitive economics are usually connected to runoff extremes. In fact, the runoff extremes are always connected to a human activity since they are not existing in a natural water cycle. Engineering science considers the runoff extremes as critical values of runoff leading to damage of infrastructure or water shortages, and introduces the extremes in terms of probability. In particular, in water engineering the runoff extremes are evaluated from tails of exceedance probability curves to be used in risk assessment of water infrastructure and decision-making in cost-lost situations (Mylne, 2002; Murphy, 1977, 1976). The exceedance probability curve (EPC) of multi-year runoff allows estimation of the runoff extremes and supports designing of building constructions, bridges, dams, withdrawal systems, *etc*.

Modern hydrology uses two approaches to evaluate the runoff extreme with their exceedance probability: conceptual modelling (Lamb, 2006) and a frequency analysis (Kite, 1977; Benson, 1968; Kritsky and Menkel, 1946). In the conceptual modelling approach, synthetic runoff series are simulated from meteorological series to calculated the runoff values of chosen exceedance probability (Arheimer and Lindström, 2015; Veijalainen et al., 2012; Seibert, 1999). In the frequency-analysis approach, historical yearly time series of runoff are used to evaluate statistical estimators, *i.e.* mean value, coefficient of variation and coefficient of skewness (van Gelder, 2006). These estimators are applied to calculate the runoff values with their exceedance probability (Guidelines SP 33-101-2003, 2004; Guidelines, 1984; Bulletin 17–B, 1982) needed to support designing of roads, dams, bridges or water-withdrawal stations. The basic assumption of this approach is that the future risks during infrastructure's operational period are equal to the risks estimated from the past observations. The runoff extremes are simply extrapolated for the next 20-30 years on an assumption that the past observations are representative for the future or a "stationarity" assumption (Madsen et al., 2013).

A number of weather extremes including hurricanes, wind, rain and snow storms, floods and droughts has increased (Vihma, 2014; Wang and Zhou, 2005, Manton et al., 2001). Historical time series of many climate variables evident trends, which are statistically significant, and the series of streamflow runoff are among others (Wagner et al., 2011; Dai et al., 2009; Milly at al., 2005). Rosmann et al. (2016) apply the Mann–Kendall Test to analyse a time series of daily, monthly and yearly river discharges for last four decades. The highest number of the trends are detected for the yearly time series of annual runoff. The statistically significant trends are founded on historical time series, thus the water engineers and managers are motivated to revise a basic "stationarity" assumption behind the infrastructures' risk assessment since the past observations are not representative for the future (Madsen et al., 2013; Kovalenko, 2009; Milly at al., 2008).

In this paper, we described a method combing the conceptual modelling and frequency analysis to estimate the runoff extremes in changing climate. The method adapts a theory of stochastic systems to a water-engineering practice, and it is further named as an Advanced of Frequency Analysis (AFA). It is introduced by Kovalenko (1993) relying on theory of stochastic systems (Pugachev et al., 1974). The basic idea behind the method is to simulate the statistical estimators of multi-year runoff (annual, minimal and maximal) from the statistical estimators of precipitation and air temperature on a climate scale (Budyko and Izrael, 1991). The simulated statistical estimators of runoff are used to construct exceedance probability curves (EPCs) with distributions from the Pearson System (Pearson, 1895). Kovalenko (1993) suggests modelling the EPCs within the Pearson Type III distribution based on a transitional practice in water engineering (Rogdestvenskiy and Chebotarev, 1974; Matalas and Wallis, 1973; Sokolovskiy, 1964). However, the distribution can be also chosen by fitting (Laio et al., 2009) or defined in accordance with local hydrological guidelines (Bulletin 17-B, 1982) or somehow more advanced (Andreev et al., 2005).

A linear "black-box" with stochastic components (or "linear filter stochastic model", LFSM) is suggested as a catchment-scale hydrological model (Kovalenko, 1993). For this linear model, the theory of stochastic systems provides methods to direct simulation of probability distributions for a random process (Pugachev et al., 1974). The theory of stochastic systems is applied to analyse and predict runoff extremes on various time scales ranging from days (Rosmann and Domínguez, 2017), to a season (Domínguez and Rivera, 2010; Shevnina, 2001) and to a climate scales (Shevnina et al., 2017; Kovalenko, 2014, Viktorova and Gromova, 2010). The AFA approach is a simplification of the theory of stochastic systems on a climate scale. Kovalenko et al., (2010) give the guidelines for water engineers on an estimation of the runoff extremes in changing climate.

The AFA has been suggested about 30 years ago, however a full description of this approach is still not published in English. Moreover, the previous publications in Russian contain many typewriting mistakes in formulas (Kovalenko, 1993; Kovalenko et al., 2006), and it makes understanding troublesome even for native Russians. In this paper, a theory and assumptions of the AFA approach were formulated "step-by-step" (in the Annex 1), and formulas behind the core of the probabilistic hydrological model MARCS^{HYDRO} were accepted for the new version 0.2 (in the Section 1). This model core allows to predict a skewness parameter of the Pearson Type III distribution. An example of the model set up, forcing and output for a case study of the Iijoki river is given in the Section 2. The main features of the model and the limitations of the AFA method were formulated in the Discussions to better place the model MARCS^{HYDRO} among other hydrological models.

1 Model description

The probabilistic hydrological model MARCS^{HYDRO} consists of six blocks Shevnina (2015). Fig. 1 shows tools for data analysis grouped into the blocks: two blocks to analysis and screening of observed data (DPB and DSB), the block with the model parametrization, cross-validation and hind casts (PHP), the block to visualize the model's results (VAB) and the block with socio-economic applications (EAB). Shevnina and Gaidukova (2017) provide details about algorithms already implemented to each block in the model. In this paper, the only version 0.2 for the model's core was introduced. The formulas behind the model's core version 0.1 is published as the annex to Shevnina et al. (2017.)

The MARCS^{HYDRO} model simulates three non-central statistical moments of multi-year runoff based on means of precipitation calculated over a period of 20–30 years. Now, the model application is limited by only a prediction on the climate scale. Development of a socio-economic infrastructure needs also for the climate scale prediction of river runoff (Milly et al., 2008) because the water extremes such as floods and droughts lead to economical losses. The AFA approach have found the practical applications to the building constructions (Shevnina et al., 2017; Kovalenko, 2009). The MARCS^{HYDRO} model allows to "quick analysis" of the runoff extremes under different climate scenarios. The model needs less computational resources because it simulates parameters of the distribution while the conceptual hydrological models simulate the runoff time series. These time series than aggregated to distributions by methods of the frequency analysis (Veijalainen et al., 2012) or with an ensemble of climate models (Madsen et al., 2013).

The MARCS^{HYDRO} model parametrization, cross-validation and hind casts needs to observations on river water discharges on a hydrological network for a period in the past. The description of the analysis and screening of the observed time series as well as the model cross-validation procedure were outside the topics of this paper. We focused on the equations behind the model's core version 0.2 and its limitations.

1.1 Model input

Two blocks of the MARCS^{HYDRO} model are needed to analysis and screening of observations (DPB and DSB). The observed time series of river runoff and precipitation are needed for the period as longer as possible. However, the length of yearly time series on water discharges usually does not exceed 80-90 years. Hydrological year books or runoff data sets provide observations at sites of National hydrological networks, and the river runoff is expressed in volumetric flow rate (water discharge, m^{3}s^{-1}). In the DPB, the volumetric rate is converted to a specific discharge (DR, mm year^{-1}):

$$DR = 1000\, Q\, T \,/\, A,$$

where Q is a yearly average water discharge, m^{3}s^{-1}; T is a number of seconds in a year, and A is the catchment area, m^2. In the DSB, the yearly time series of DR are used to analyze for homogeneity and trends (Dalmeh and Hall, 1990) and to define a period for the model parametrization, or "a reference period" in Shevnina at al. (2017). Then, the reference three non-central moments m_k ($m_k = 1/n \sum_{i=1}^{n} DR_i^k$ for k=1, 2, 3) are estimated from time series of DR with a Method of Moments (van Gelder et al., 2006).

The observations on a precipitation are collected on meteorological sites, and they may be interpolate into grids to better estimate a precipitation rate over a river basin area. In the DPB, the mean annual precipitation rate (AP, mm year^{-1}) is calculated from the observed yearly time series for the reference period. The mean of AP for the future period can be calculated from an output of any global/regional climate model or even a set of models. In a study on a catchment scale, the time series of water discharges can be extracted from the Global Runoff Data Center (GRDC) while the precipitation rate can be estimated from gridded data sets (Willmott and Robeson, 1995). These two data sets were used to perform an example of the model application the Iijoki River river basin.

1.2 Model cross-validation

MARCS^{HYDRO} allows to simulate the non-central moments of runoff to be used for construction of probability distribution (or exceedance probability curve), *i.e.* provides a probabilistic form of prediction. The end product of the model is the PDF (or EPC), and there are no simulated time series of runoff to be compared with observations. Kovalenko (1993) suggests to compare the simulated PDF with empirical PDF by known statistical tests such as the Kolmogorov-Smirnov test (Smirnov, 1948). In the PHB of the MARCS^{HYDRO} model, a specific cross-validation procedure allows conclusions about the model's validation and quality of hind casts. For the model's cross validation, the observed time series of river runoff is divide into two sub periods namely training and control. The splitting year is corresponded to a year when a statistically significant difference of mean values estimated over two periods. In this study, we did not pay much attention to the cross-validation procedure described in Shevnina et al. (2017) and Kovalenko (1993).

1.3 Model core

In our study, the core version 0.2 for the probabilistic model MARCS^{HYDRO} was suggested instead of the version 0.1 (Shevnina et al., 2017). The version 0.2 allows an evaluation of a skewness parameter of the Pearson Type III distribution. In the new core, the non-central statistical moments of the *DR* were calculated as following:

$$m_1 = a - b_1 , \tag{1}$$

$$m_2 = -b_0 - 2 m_1 b_1 + m_1 a , \tag{2}$$

$$m_3 = -2 m_1 b_0 - 3 m_2 b_1 + m_2 a , \tag{3}$$

where, m_1, m_2 and m_3 are the moment estimates of the non-central statistical moments of the *ARR*; a, b_0, b_1 and b_2 are the parameters of distribution from the Pearson System (Andreev et al., 2005) denoted as PSD in the further text.

To set up the MARCS^{HYDRO} model the observations on water discharges are needed. For the reference period (notated by low index r) the moments' estimates for the non-central moments (m_{1r}, m_{2r}, m_{3r}) were calculated from observed times series of runoff (mm year^{-1}) first. Then, the non-central moments were used to evaluate the parameters of the Pearson equation a, b_0, b_1 :

$$a = 0.5\left(5 m_{1r} m_{2r} - 4 m_{1r}^3 - m_{3r}\right) / \left(m_{2r} - m_{1r}^2\right) , \tag{4}$$

$$b_0 = 0.5\left(m_{1r}^2 m_{2r} - 2 m_{2r}^2 + m_{1r} m_{3r}\right) / \left(m_{2r} - m_{1r}^2\right) , \tag{5}$$

$$b_1 = 0.5\left(3 m_{1r} m_{2r} - 2 m_{1r}^3 - m_{3r}\right) / \left(m_{2r} - m_{1r}^2\right) . \tag{6}$$

Then, the parameters of the linear filter model (LFSM, see the Annex 1 for details) $\bar{c}, G_{\tilde{N}}, G_{\tilde{c}\tilde{N}}$ denoted by low index r were calculated:

$$\bar{c}_r = \bar{N}_r / \left(a - b_1/2\right) , \tag{7}$$

$$G_{\tilde{N}r} = -2 b_0 \bar{N}_r / \left(a - b_1/2\right) , \tag{8}$$

$$G_{\tilde{c}\tilde{N}r} = b_1 \bar{N}_r / \left(a - b_1/2\right) , \tag{9}$$

where, $\bar{N}_r$ is the mean of *AP* (mm year^{-1}) estimated from observed time series as an average over any chosen reference period.

To force the MARCS^{HYDRO} model the outputs from global/regional scale climate models are needed. The CMIP5 (Taylor et al., 2012) is among other collections of data sets available for climate scale hydrological studies. Recently, the model needs to be forced only a mean of precipitation ($\bar{N}$, mm year^{-1}) evaluated of the future period of 20-30 year. The low index *pr* indicated that the values were estimated for the future, and the $\bar{N}_{pr}$ are estimated from climate scenarios. In an assumption that $\bar{c}, G_{\tilde{N}}, G_{\tilde{c}\tilde{N}}$ are constant for both periods $\bar{c}_r=\bar{c}_{pr}, G_{\tilde{N}r}=G_{\tilde{N}pr}, G_{\tilde{c}\tilde{N}r}=G_{\tilde{c}\tilde{N}pr}$ (a "basic parametrization scheme" according to Kovalenko, 1993), new parameters of the PSD are calculated from the $\bar{N}_{pr}$:

$$a=\left(G_{\tilde{c}\tilde{N}pr}+2\bar{N}_{pr}\right)/\left(2\bar{c}_{pr}\right) , \tag{10}$$

$$b_0=-G_{\tilde{N}pr}/\left(2\bar{c}_{pr}\right) , \tag{11}$$

$$b_1=G_{\tilde{c}\tilde{N}pr}/\bar{c}_{pr} , \tag{12}$$

Finally, the non-central moments of runoff are calculated for the projected period (denoted by low index *pr*):

$$m_{1pr}=a-b_1 , \tag{13}$$

$$m_{2pr}=-b_0-2m_{1pr}b_1+am_{1pr} , \tag{14}$$

$$m_{3pr}=-2m_{1pr}b_0-3m_{2pr}b_1+am_{2pr} . \tag{15}$$

It should be noted, that in the core version 0.2, the linear filter model includes the multiplicative stochastic component (see the Annex 1 for details). It may leads to unstable solutions of the Fokker-Plank-Kolmogorov (FPK) equation ($m_k \rightarrow \infty$) for the statistical moments of high orders. Two methods to get stable FPK solutions are suggested by Kovalenko (2004), and one of them is already implemented in the core version 0.1.

1.4 Model output

In our study, the exceedance probability curve (EPC) of runoff was modelled within the Pearson Type III distribution. This distributions is commonly used by water engineers to estimate water extremes (Kountrouvelis and Canavos, 1999; Rogdestvenskiy and Chebotarev, 1974; Matalas and Wallis, 1973). The water engineering guidelines provide the ordinates of EPCs from look-up tables (Guidelines, 1984) depending on a coefficient of variation (CV) and coefficient of skewness (CS). These coefficients are calculated from non-central moments' estimates (Rogdestvenskiy and Chebotarev, 1974):

$$\mathrm{CV}=\sqrt{\left(m_2-m_1^2\right)}/m_1 , \tag{16}$$

$$\mathrm{CS}=\left(m_3-3m_2m_1+2m_1^3\right)/\mathrm{CV}^3m_1^3 . \tag{17}$$

The the MARCS$^{\mathrm{HYDRO}}$ model output includes the estimates of the mean value, CV and CS calculated for the reference period from observations as well as these estimates simulated from mean precipitation for the projected period. The ordinates of EPC available from look-up tables allows then to calculate the runoff values together with their exceedance probability.

2 Model application: a case study

In our study, we chosen the river basin of the Iijoki at Raasakka gauge (Lat 25.411º / Lon 65.335º) to give an example of the application of the MARCS^{HYDRO} model on a catchment scale. The Iijoki river is located north west Finland, and the Raasakka gauge outlines the watershed area of over 14,191 km^2. The catchment has a small population and there are no hydro power plants of multi-year regulation to affect natural regime on annual cycle. Thus, one can expect that historical yearly time series of annual runoff rate do not contain trends connected to the artificial regulation. This case study shows an example of the set up and output of the probabilistic model MARCS^{HYDRO}.

2.1 Model set up: the reference period

The yearly time series of volumetric water discharge of the Iijoki river were extracted from a dataset of the Global Runoff Data Center (GRDC, 56068 Koblenz, Germany). The observations at the Raasakka gauge (ID = 6854600) cover a period 1911–2014, and they do not contain gaps. This period was considered as the reference. The annual specific discharge (DR, mm year^{-1}) was calculated from the average volumetric water discharge for each year in the reference period. Then, the non-central moments were calculated from the yearly time series of DR with the Method of Moments (Table 1). The reference climatology (the means of precipitation and air temperature) were evaluated from the dataset of NOAA (NOAA/OAR/ESRL PSD, Boulder, Colorado, USA) at a grid node nearest to the watershed centroid (this technique will be discussed in a separate paper as well as the methods of a forcing pre-analysis).

2.2 Model forcing: the projected period

Climate scenarios provide a range of projections for temperature and moisture regimes in the future. This range is produced by different assumptions behind climate scenarios as well as a specific of climate models. However, the climate projections include precipitation and air temperature, and they give a forcing to hydrological models to simulate projections of runoff. In the case study of the Iijoki river, the data from the Coupled Model Inter-comparison Project 5, CMIP5 (Taylor et al., 2012) for three Representative Concentration Pathways (RCPs) were used to force the MARCS^{HYDRO} model. For each RCP scenario, the projections of annual precipitation rate were applied to test how the MARCS^{HYDRO} model simulates the EPD under different forcing trajectories. For the period of 2020–2050 (considered as the projected), the mean values of precipitation rate ($\bar{N}_{pr}$, mm year^{-1}) were calculated based on four world-leading global climate models. We used the outputs from the CaESM2 (Chylek et al., 2011), HadGEM2-ES (Collins et al., 2011), INM-CM4 (Volodin et al., 2010) and MPI-ESM-LR (Giorgetta et al., 2013) global models (Table 2). The $\bar{N}_{pr}$ varied by 2–5 % of the model's average over the RCP scenarios, however, these values alter substantially between the climate models. Among the outputs considered, the MPI-ESM-LR model projects highest changes in the $\bar{N}_{pr}$ compared to the reference period (Tables 1 and 2). The

HadGEM2-ES model gives the lowest values for the $\bar{N}_{pr}$. The projected means of precipitation rate are slightly varied between the scenarios. At the same times, they are exhibit the significant range of changes among the climate models (the $\bar{N}_{pr}$ range from 619 to 737 mm year^{-1}) for the case of the Iijoki river at Raasaka.

2.3 Model output: projected period

The projected non-central moments' estimates were simulated for the scenarios/models listed in Table 2. These estimates were used to calculate the mean value, CV and CS (see the Eq. 16–17) included to the output of the MARCS^{HYDRO} model. Table 3 shows the modelling results for the HadGEM2-ES and MPI-ESM-LR global models, were the water discharges of 10 and 90 % exceedance probabilities are given. The ordinates of the Pearson Type III distribution were extracted from the look-up tables used in hydrological engineering (Druzhinin and Sikan, 2001), and they allow to expressed runoff as volumetric rate or water dicharge (m^{3}s^{-1}). For the Iijoki River at Raasakka, the mean values of *DR* and CV vary of over 7 % and 5 % correspondingly under the RCP scenarios. The maximum alteration in the projected mean values of *DR* were obtained under RCP85 (619 to 737 mm year^{-1}). Under the projections of the MPI-ESM-LR model, the mean of *DR* increases of over 17 %.

In the case of the Iijoki River, the water discharge of 10 % exceedance probability are going to increase in the future under scenarios/models considered (Table 3). It may leads to risks of energy spills at hydropower stations during the period 2020–2050. At the same time, risks connected to water shortage may be less since they are connected to water discharge of 90 % exceedance probability which are predicted to increase. Figure 2 shows another form how the model performs the EPC of annual runoff rate for the Kyrönjoki River at Skatila (GRDC ID: 6854900). The set of EPCs were simulated under three RCP scenarios using a similar set up of the MARCS^{HYDRO} model (will be discussed in a separate paper). In further development of the visualisation block, it would important to involve water managers and decision makers to better outline practical applications for the probabilistic hydrological model.

Discussions

Nowadays, a future vision of the climate is changing continuously. The climate projections are updated almost every 5–6 years and many climate models generate meteorological projections for variables such as precipitation and air temperature. It needs to hydrological models to perform an "express analysis" about future changes in water resources and water extremes (floods and droughts) on a climate scale. The climate scale means that the "express analysis" is provided for the period of 20-30 years. The lumped or semi-distributed physically-based hydrological models traditionally used for short term or seasonal scale simulating runoff times series from time series of meteorological variables (Seibert, 1999). In many catchment scale hydrological studies these models are driven by outputs of climate models or their ensemble to evaluate water resources and extremes in the near future (Arheimer and Lindström, 2015; Veijalainen et al., 2012; Yip et al., 2012). The simulation of the runoff time series from the time series of meteorological variables (see Fig. 2 in Veijalainen et al. (2012)) leads to high computational costs of such estimations needed to be served in term of probability in economical applications (Murphy, 1976). The probabilistic MARCS^{HYDRO} model is computationally cheaper while to compare to lumped or semi-distributed physically-based hydrological models. It can be easy decoupled with global and regional climate models and to provide the "express analysis" of water resources under a modern version of the future climate.

In this paper we described the structure for the probabilistic hydrological model MARCS^{HYDRO} together with the AFA method behind a new model's core version 0.2. The AFA method has more than 20 years story, however most of studies is published in Russian (Kovalenko, 1993; 2004; 2009; Kovalenko et al., 2010). The AFA method is based on the statistical theory of automatic system (Pugechev et al., 1974) , which is an outsider among the "classical hydrological" disciplines. The AFA method is one of simplification of the Fokker-Plank-Kolmogorov equation approach been developed in the Russian State Hydrometeorological University. It is tested in many case studies on river basins located in Russia, Colombia, Bolivia, Mali, *etc*.. There are also a number of publications in English (Rosmann and Domínguez, 2017; Shevnina et al., 2017; Kovalenko, 2014; Domínguez and Rivera, 2010; Viktorova and Gromova, 2008). In this manuscript we formulated the theory logically in an attempt to provide the equations for the new core 0.2 of the model MARCS^{HYDRO} model, however it needs to describe also the AFA method behind.

The probabilistic hydrological model MARCS^{HYDRO} includes the core versions 0.1 and 0.2. In both cores, the only three non-central moments are evaluated to construct the exceedance probability curve within the theoretical distribution the Pearson III type, which is among traditional distributions on the frequency and risk analysis in hydrology (Kite, 1977; Rogdestvenskiy and Chebotarev, 1974; Sokolovskiy, 1968; Elderton, 1969; Benson, 1968). The model simulates three estimates of non-central moments of runoff instead of the runoff time series, and this circumstance makes the computations by the MARCS^{HYDRO} model to be a "low cost" compared to conceptual hydrological models (Arheimer and Lindström, 2015; Veijalainen et al., 2012). The MARCS^{HYDRO} model also allows to put the projections of runoff in term of probability, *i.e.* as runoff values together with their exceedance probability.

The MARCS^{HYDRO} model includes six modules, and each module allows improvements by including a new methods. In this paper, the new model core version 0.2 extending to simulate the third statistical estimator (skewness) was presented. The applicability of the core version 0.2 is limited by assumptions behind the AFA approach. The "quasi-stationary" assumption for the expected climate change is among others. In this case, the climate is described by statistical estimators *i.e.* mean value, variability, *etc.* of precipitation, air temperature, evapotranspiration, river runoff *etc.* for the period of 20–30 year. It is assumed to consider two time period periods with statistically different climate namely the reference and projected periods. Another limitation is connected to the linear filter stochastic model (see details in the Annex 1) used in the core version 0.2. It should be noted that there is a multiplicative component in the model core, and it may lead to unstable solutions of the FPK equation. Kovalenko (2004) suggests two solutions resulting to the stable solutions of the FPK. On of the solution is given by Kovalenko et al. (2010) and is coded in the model version 0.1 (Shevnina et al., 2017). However, a checking procedure needs to be apply to before using this core version. In the checking procedure we plan to use a "beta criterion" method suggested in Kovalenko (2004) to develop the MARCS^{HYDRO} model.

Further improvements of the MARCS^{HYDRO} model are going to be further implemented in the block of parametrization and hind casts. Recently the only basic parametrization scheme (Kovalenko, 1993) is included. This basic scheme gives over 70–80 % successful hind casts ("forecasts in the past") in the model cross-validation (Shevnina et al., 2017), and the implementation of the regional oriented parametrization scheme (Shevnina, 2011) is our next step further. It needs to include a mean value of air temperature to the parameter connected to "noised" watershed physiography in Eq. (A.4), the inverse of runoff coefficient in Kovalenko, (1993). It is also important to study a role of spatial resolution of meteorological forcing to affect the modelling uncertainties for simulated mean, CV and CS of runoff.

To fine the probabilistic MARCS^{HYDRO} model among other hydrological models, its practical applications needs to be better outlined. The model serves a probabilistic form of long-term hydrological projections, and they require to be adapted for needs of water engineers and water managers as a tool for risks analysis under expected climate change. The projected exceedance probability curves of multi-year river runoff can be applied in designing of bridges, pipes, dams *etc.* to minimize the future risks connected to extreme floods (Shevnina et al., 2017; Kovalenko et al., 2014; Kovalenko, 2009) or to water shortage due to droughts (Viktorova and Gromova, 2014). It is important to define informative forms for the outputs of the model MARCS^{HYDRO} to be adapted for needs of a practice, and the development of the block of economic application is among others studies to be continued in close cooperation with water managers and decision makers.

Conclusion

The paper describes the theory and assumptions of the AFA approach as well as the probabilistic hydrological model MARCS^{HYDRO} structure and core version 0.2. The features of the model are: the close connection to water engineering due to serving the runoff projection in terms of probability, cheapness in term of computational cost and a wide range of techniques allowing the model improvement. In the new core, the third moment linked to the location parameter of the Pearson Type III distribution (or asymmetry) was implemented to be simulated. In the previous version of the model core, a constant ratio CS/CV is used to calculate the location parameter of the distribution.

To give a practical example how to set up the MARCS^{HYDRO} model, the case of the Iijoki River at Raasakka (Finland) was considered. The model simulated the tailed values of 10 % an 90 % of annual runoff from the outputs of global climate models. We shown two forms of the probabilistic projections of runoff: as the exceedance probability curve and as the runoff values with their exceedance probability. This case study of the Iijoki River at Raasakka shows that the MARCS^{HYDRO} model gives reasonable results for the meteorological projections considered. The practical applications of water management and decision making should be clarified in further studies in close co-operation with water engineers.

Code availability

Currently, the MARCS^{HYDRO} model code is hosted in: https://github.com/ElenaShe000/MARCS with details on its applications for catchment scale case studies. The model source code for the core version 0.2 is distributed under the Creative Commons Attribution 4.0 License and can be downloaded from the link: https://zenodo.org/record/1220096#.WyTXxxxRVhw, and used freely in a scientific research with reference to this publication. We hope that this type of license provides the best way to create a community of motivated people to further development the model. Then, the source code will be distributed under the terms of a user agreement.

Data availability

The following data sets can be used to set up and forcing the MARCS model: the Global Runoff Data Center (GRDC, 56068 Koblenz, Germany), the NOAA/OAR/ESRL PSD (Boulder, Colorado, USA) as well as the Coupled Model Inter-comparison Project 5, CMIP5 (Taylor et al., 2012).

Sample availability

The sample data set for the Iijoki River at Raasaka case study is given in https://zenodo.org/record/1220096#.WyTXxxxRVhw.

Annex 1. Theoretical basis for the core version 0.2

A1.1 Assumptions behind the Advance of Frequency Analysis (AFA)

The Advance of Frequency Analysis is based on the theory of stochastic systems, specifically, the Fokker-Plank-Kolmogorov equation (FPK), which is simplified to a system for three non-central statistical moments (Pugachev et al., 1974). The time series of annual runoff is considered as realization of a random process Markov chain type assumed to be "stationary". It means that the statistical estimators (mean, variance and skewness) do not change over period considered. The statistical estimators are used to model an exceedance probability curve of annual runoff within the Pearson Type III distribution. The AFA approach is developed with an assumption of "quasi-stationary" (Kovalenko et al., 2010, Kovalenko, 1993). The "quasi-stationary" assumption suggests that the statistical estimators of multi-year runoff are different for two periods (reference and projected). For the reference period, the statistical estimators are evaluated from historical yearly time series of runoff. For the projected period, the statistical estimators of runoff are simulated based on an output of global- or regional-scale climate models under any climate scenario.

A1.2 Linear filter stochastic model (LFSM)

Models replace a complicate hydrological system by maths abstractions, and aim to reveal spatial and temporal runoff features which are important depending on goals of study. Among others, “black box” hydrological models consider a river basin as a dynamical system with lumped parameters. These models are “based on analysis of concurrent inputs and temporal output series” (WMO-№168, 2009), and transform series of meteorological variables (precipitation, air temperature) into series of runoff. Both input and output series are functions of time (WMO-№168, 2009):

$$a_n(t)\frac{d^nQ}{dt^n}+a_{n-1}(t)\frac{d^{n-1}Q}{dt^{n-1}}+...+a_1(t)\frac{dQ}{dt}+a_0(t)Q= \\ =b_n(t)\frac{d^nP}{dt^n}+b_{n-1}(t)\frac{d^{n-1}P}{dt^{n-1}}+...+b_1(t)\frac{dP}{dt}+b_0(t)P \quad , \tag{A.1}$$

where Q is the runoff in volumetric flow rate, P is the precipitation in volumetric flow rate (rain, snow melt); and the coefficient a_i and b_i are the empirical parameters of a translating system. These coefficients are lumped parameters of the “black box” model. The solution to Eq. (A.1) for zero initial conditions gives (WMO-№168, 2009):

$$Q(t)=\int_0^t h(t,\tau)P(\tau)d\tau \ , \tag{A.2}$$

where the function $h(t,\tau)$ represents a response of a river basin at time *t* to a single portion of precipitation at time τ. In the AFA approach, a river basin is considered as a linear system transforming the annual precipitation into the annual runoff:

$$a_1(t)\frac{dQ}{dt} + a_0(t)Q = b_0(t)P \ . \tag{A.3}$$

On the other hand, a river basin can be considered as a linear system with stochastic components in the input function and the model parameter:

$$dQ=[-(\bar{c} + \tilde{c}(t))Q + (\bar{N} + \tilde{N}(t))]dt \ , \tag{A.4}$$

where $a_0(t)=\bar{c}+\tilde{c}(t)$ is the stochastic parameter of the system (a “noised” watershed physiography, the inverse of runoff coefficient in Kovalenko, (1993)); $b_0(t)P=\bar{N}+\tilde{N}(t)$ is the stochastic input for the system (a “noised” precipitation), and $a_1=1$. The stochastic components of $\tilde{c}(t)$ and $\tilde{N}(t)$ are the Gaussian “white noise” with zero means, and their intensities are $G_{\tilde{c}}$, $G_{\tilde{N}}$. The intensities are mutually correlated as $K_{\tilde{c}\tilde{N}}(\tau)=E(\tilde{c}(t)\tilde{N}(t+\tau))=G_{\tilde{c}\tilde{N}}\delta(\tau)$. It should be noted, that the multiplicative parameter $\bar{c} + \tilde{c}(t)$ in the Eq. (A.4) is the sum of constant $\bar{c}$ and Gaussian «white noise» $\tilde{c}(t)$, and it may lead to the unstable solutions of the FPK equation (to infinite of statistical moments of high orders). It limits application of the AFA method (Kovalenko, 1993). Kovalenko (2004) suggests two solutions, and we will introduce them in a further paper.

A1.3 Fokker-Plank-Kolmogorov equation (FPK) and simplifications

The Fokker-Plank-Kolmogorov (FPK) equation can be applied to simulate the probability density function (PDF) for the stochastic *Q(t)* in Eq. 4 (Kovalenko, 1993; Pugachev, 1974):

$$\frac{\partial p(Q,t)}{\partial t} = -\frac{\partial}{\partial Q}\left(A(Q)\,p(Q,t)\right) + 0.5\frac{\partial^2}{\partial Q^2}\left(B(Q)\,p(Q,t)\right) , \qquad \text{(A.5)}$$

where $p(Q,t)$ is the PDF of Q at time t; and the drift coefficient ($A(Q)$) and diffusion coefficients ($B(Q)$) are calculated as follows (Kovalenko, 1993; Pugachev, 1974):

$$A(Q) = -\left(\bar{c} - 0.5G_{\tilde{c}}\right)Q - 0.5G_{\tilde{c}\tilde{N}} + \bar{N} , \qquad \text{(A.6)}$$

$$B(Q) = G_{\tilde{c}}Q^2 - 2QG_{\tilde{c}\tilde{N}} + G_{\tilde{N}} . \qquad \text{(A.7)}$$

The analytical solution of Eq. (A.5) is difficult and not always needed for practical applications in water engineering since the PDFs of runoff are modelled from a set of statistical estimators, and moments are from, among others, van Gelder et al. (2006). The PDFs are described with the set of moments $m_k = \int_{-\infty}^{+\infty} Q^k p(Q,t)\,dQ$ (where *k* is number of the moment, *k* –> ∞). To obtain the equations for m_k , both sides of Eq. (A.5) were multiplied by a differentiable function $\psi(Y)$ and then were integrated within limits from $-\infty$ to $+\infty$ by Q (however, it is supposed that $Q > 0$):

$$\frac{d\left(\int_{-\infty}^{+\infty}\psi(Q)\,p(Q,t)\,dQ\right)}{dt} = \int_{-\infty}^{+\infty} p(Q,t)\,A(Q)\frac{\partial\psi(Q)}{\partial Q}dQ + 0.5\int_{-\infty}^{+\infty} p(Q,t)\,B(Q)\frac{\partial^2\psi(Q)}{\partial Q^2}dQ \qquad \text{(A.8).}$$

Then, $\psi(Q)$ was replaced with $\psi(Q) = Q^k$, and the Eq. (A.8) was written as:

$$\frac{dm_k(t)}{dt} = \int_{-\infty}^{+\infty} p(Q,t)\,A(Q)\frac{\partial\left(Q^k\right)}{\partial Q}dQ + 0.5\int_{-\infty}^{+\infty} p(Q,t)\,B(Q)\frac{\partial^2\left(Q^k\right)}{\partial Q^2}dQ . \qquad \text{(A.9)}$$

For a stationary random process $dm_k(t)/dt = 0$, and the drift and diffusion coefficients are constant. Thus, Eq. (A.9) was simplified as follows:

For *k*=1:

$$-\left(\bar{c} - 0.5G_{\tilde{c}}\right)m_1 - 0.5G_{\tilde{c}\tilde{N}} + \bar{N} = 0 . \qquad \text{(A.10)}$$

For $k \geq 2$:

$$-k\left(\bar{c}-0.5kG_{\tilde{c}}\right)m_k+ k\bar{N}m_{k-1}-k\left(k-0.5\right)G_{\tilde{c}\tilde{N}}m_{k-1}+0.5k\left(k-1\right)G_{\tilde{N}}m_{k-2}=0 \ . \qquad \text{(A.11)}$$

Further, the summands in Eq. (10–11) were divided by $\left(2\bar{c}+G_{\tilde{c}}\right)$, and new notations were introduced as suggested in (Kovalenko, 1993; Pugachev et al., 1974):

$$a=\frac{G_{\tilde{c}\tilde{N}}+2\bar{N}}{2\bar{c}+G_{\tilde{c}}} \ ; \ b_0=-\frac{G_{\tilde{N}}}{2\bar{c}+G_{\tilde{c}}} \ ; \ b_1=\frac{2G_{\tilde{c}\tilde{N}}}{2\bar{c}+G_{\tilde{c}}} \ ; \ b_2=-\frac{G_{\tilde{c}}}{2\bar{c}+G_{\tilde{c}}} \ .$$

Then, for k = 1, 2, 3, 4 the system of Eq. (A.10–11) includes:

$$m_1\left(2b_2+1\right)-a+b_1=0 \ , \qquad \text{(A.12)}$$

$$\left(3b_2+1\right)m_2+\left(2b_1-a\right)m_1+b_0=0 \ , \qquad \text{(A.13)}$$

$$\left(4b_2+1\right)m_3+\left(3b_1-a\right)m_2+2b_0m_1=0 \ , \qquad \text{(A.14)}$$

$$\left(5b_2+1\right)m_4+\left(4b_1-a\right)m_3+3b_0m_2=0 \ . \qquad \text{(A.15)}$$

The set of four moments (m_1, m_2, m_3, m_4) is sufficient to model distributions from the Pearson System (Andreev et al., 2005; Elderton and Johnson, 1969). However, in water engineering we usually use only three-parametric probability distributions fitted to observations (Guidelines, 2004; Guidelines, 1984; Bulletin 17-B, 1982). In this case, the $G_{\tilde{c}} \ll \bar{c}$ is assumed, thus it leads to $b_2=-G_{\tilde{c}}/\left(2\bar{c}+G_{\tilde{c}}\right)\approx 0$ and $\left(4b_2+1\right)\approx 1$, $\left(3b_2+1\right)\approx 1$, $\left(2b_2+1\right)\approx 1$. To model the PDFs (or EPCs) of annual runoff within the Pearson Type III distribution, the system of Eq. (A.12–15) is simplified as follows:

$$-a+b_1=-m_1 \ , \qquad \text{(A.16)}$$

$$b_0+2m_1b_1-am_1=-m_2 \ , \qquad \text{(A.17)}$$

$$2m_1b_0+3m_2b_1-am_2=-m_3 \ . \qquad \text{(A.18)}$$

Denoting $lk=\begin{pmatrix}-m_1\\-m_2\\-m_3\end{pmatrix}$, $x=\begin{pmatrix}b_1\\b_0\\a\end{pmatrix}$ and $A=\begin{pmatrix}1 & 0 & -1\\ 2m_1 & 1 & -m_1\\ 3m_2 & 2m_1 & -m_2\end{pmatrix}$, the parameters a, b_0, b_1 are calculated as $x_i = D_i / D$, where D is the determinant of matrix A, and D_i is the determinant of the matrix obtained by replacing of the column i (1, 2, 3) in matrix A by the vector lk. Finally, the parameters a, b_0, b_1 are calculated as follows:

$$b_1=0.5\left(3m_1m_2-2m_1^3-m_3\right)/\left(m_2-m_1^2\right) \ , \qquad \text{(A.19)}$$

$$b_0=0.5\left(m_1^2m_2-2m_2^2+m_1m_3\right)/\left(m_2-m_1^2\right) \ , \qquad \text{(A.20)}$$

$$a=0.5\left(5m_1m_2-4m_1^3-m_3\right)/\left(m_2-m_1^2\right) \ . \qquad \text{(A.21)}$$

A1.3 Notations

There are too many notations used while to describe the model core version 0.2, thus the secondary parameters of equations were grouped by model behind. Table A.1 shows the notation and description of the secondary parameters for the linear filter stochastic model. The Eq. A.3 is simplification of the Eq. A.1 by limiting first order ordinal differential equation. It includes three parameters a_0 , a_1 and b_0 , and two of them are assumed to be “noised”. These “noised” parameters include a constant component (with bar) and Gaussian “white noise” component (with tilde) with own intensities.

Table A.1 The notation and description of the parameters for a linear filter stochastic model.

Q	runoff volumetric flow rate, m^3s^{-1}
P	precipitation volumetric flow rate, m^3s^{-1}
$a_i(t)$, $b_i(t)$	lumped parameters of “block box” model, i = 0 and 1
$\bar{c}+\tilde{c}(t)$	inverse of runoff coefficient: $\bar{c}$ is constant component, $\tilde{c}(t)$ is the Gaussian “white noise”
$\bar{N}+\tilde{N}(t)$	Precipitation: $\bar{N}$ is constant component, $\tilde{N}(t)$ is the Gaussian “white noise”
$G_{\tilde{c}}$, $G_{\tilde{N}}$	intensities of the Gaussian “white noise”
$G_{\tilde{c}\tilde{N}}\delta(\tau)$	Correlation function for the mutually delta-correlated processes $\tilde{c}(t)$ and $\tilde{N}(t)$

Table A.2 gives description of the parameters of the FPK and the Pearson System Distribution (PSD). It should be noted that we do not solve the FPK, and only its simplification to the system for three non-central moments is applied. These non-central moments are estimated from runoff observations for the reference period. For the projected period the moments are calculated from a mean of precipitation.

Table A.2. The notations of the FPK equation and PSD

$p(Q,t)$	probability density function of Q at time t
$A(Q)$	drift coefficient (the FPK), estimated from the “noised” parameters and their intensities
$B(Q)$	diffusion coefficient (the FPK), estimated from the “noised” parameters and their intensities
m_k	non-central statistical moment with order k = 1, 2, 3, 4
a, b_0, b_1,b_2	parameters of a distribution (the PSD)

Annex 2. Short user guide the MARCS model

To set up the model for a single river catchment, the non-central moments should be calculated from historical time series of annual river runoff rate as well as a mean value of annual precipitation rate. These values should be placed manually (lines 45-48 in model_core.py located in https://zenodo.org/record/1220096#.WyTXxxxRVhw) as well as the ID number of catchment (line 51, model_core.py). To force the model, the projected mean value of annual precipitation rate should be evaluated from an output of climate model, and then the model_core.py can be running in Unix command line: ./model_core.py XXX (where XXX is the mean of annual precipitation rate for the projected period). The output of the model_core.py in stored in the output file model_GPSCH.txt and include line with following format: the ID of catchment, the first non-central moment estimate of annual runoff rate (mm year^{-1}) for a reference period, the mean value of annual precipitation rate (mm year^{-1}) for a reference period, the coefficient of variation for a reference period, the coefficient of skewness for a reference period, the model parameters $\bar{c}$, $G_{\tilde{N}}$, $G_{\tilde{c}\tilde{N}}$, the the first non-central moment estimate of annual runoff rate (mm year^{-1}) for a projected period, the mean value of annual precipitation rate (mm year^{-1}) for a projected period, the coefficient of variation for a projected period, the coefficient of skewness for a projected period.

Competing interests

The authors declare that they have no conflict of interests.

Acknowledgement

The study is supported by the Academy of Finland (contract 283101).

References

Andreev, A., Kanto, A. and Malo, P.: Simple approach for distribution selection in the Pearson System, working papers / Helsingin kauppakorkeakoulu, Finland, 25 pp., 2005.

Arheimer, B. and Lindström, G.: Climate impact on floods: changes in high flows in Sweden in the past and the future (1911–2100), Hydrol. Earth Syst. Sci., 19, 771-784, doi:10.5194/hess-19-771-2015, 2015.

Benson, M. A.: Uniform Flood Frequency Estimating Methods for Federal Agencies, Water Resour. Res., 4(5), 891–908, doi:10.1029/WR004i005p00891, 1968.

Budyko, M. I. and Izrael, Y. A. (Eds.): Anthropogenic Climatic Change, University of Arizona Press, Tucson, USA, 1991.

Bulletin 17–B: Guideline for determining flood flow frequency, U.S. Geological Survey, Virginia, USA, 1982.

Chylek, P., Li, J., Dubey, M. K., Wang, M., and Lesins, G.: Observed and model simulated 20th century Arctic temperature variability: Canadian Earth System Model CanESM2. Atmos. Chem. Phys. Discuss., 11, 22893–22907, doi:10.5194/acpd-11-22893-2011, 2011.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C. D., Joshi, M., Liddicoat, S., Martin, G., O'Connor, F., Rae, J., Senior, C., Sitch, S., Totterdell, I., Wiltshire, A., and Woodward, S.: Development and evaluation of an Earth-system model – HadGEM2, Geosci. Model Dev. Discuss., 4, 997–1062, doi:10.5194/gmdd-4-997-2011, 2011.

Dahmen, E.R. and Hall, M.J.: Screening of hydrological data: tests for stationarity and relative consistence, International institute for Land reclamation and improvement, The Netherlands, 1990.

Dai, A., Qian, T., Trenberth, K.E. and Milliman, J.D.: Changes in continental freshwater discharge from 1948 to 2004, J. Clim., 22 (10), 2773–2792, doi: 10.1175/2008JCLI2592.1, 2009.

Domínguez, E. and Rivera, H.: A Fokker–Planck–Kolmogorov equation approach for the monthly affluence forecast of Betania hydropower reservoir, J. Hydroinform., 12(4), 486–501, doi: 10.2166/hydro.2010.083, 2010.

Druzhinin, V. S. and Sikan A.V.: Methods of statistical processing of hydrometeorological information, Russian State Hydrometeorological University Press, St. Petersburg, 2001. (In Russian).

Elderton, S.W.P. and Johnson, N.L.: Systems of Frequency Curves, Cambridge University Press, London, 1969.

Giorgetta, M., Jungclaus, J., Reick, C., Legutke, S., Bader, J., Böttinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.-D., Ilyina, T., Kinne, S., Kornblueh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K., Claussen, M., Marotzke, J. and Stevens, B.: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, J Adv Model Earth Sy, 5(3), 572–597, doi:10.1002/jame.20038, 2013.

Guidelines to estimate basic hydrological characteristics, Gidrometeoizdat, Leningrad, USSR, 1984. (In Russian).

Guidelines SP 33-101-2003: to estimate basic hydrological characteristics, Gosstroy, Moscow, Russia, 2004. (In Russian).

Kite, G.W.: Frequency and risk analysis in hydrology, Fort Collins, Colorado, 1977.

Koutrouvelis, I. A. and Canavos, G. C.: Estimation in the Pearson type 3 distribution, Water Resour. Res., 35(9), 2693–2704, doi:10.1029/1999WR900174, 1999.

Kovalenko, V. V.: Modelling of hydrological processes, Gidrometeizdat, St. Petersburg, Russia, 1993. (In Russian).

Kovalenko, V. V.: Partial infinite modelling and forecasting of the process of river-runoff formation. - St. Petersburg: RSHU Publishers, 2004. Available on-line: http://elib.rshu.ru/files_books/pdf/img-417153826.pdf

Kovalenko, V.V.: Hydrological security of building projects with climate change, Russian State Hydrometeorological University Press, St. Petersburg, Russia, 2009. (In Russian).

Kovalenko, V. V., Victorova, N. V., Gaydukova, E. V., Gromova, M. A., Khaustov, V. A. and Shevnina, E. V.: Guideline to estimate a multi-year runoff regime under non-steady climate to design hydraulic contractions, Russian State Hydrometeorological University Press, St. Petersburg, 2010. (In Russian).

Kovalenko, V. V., Victorova, N. V., Gaydukova, E. V.: Modelling of hydrological processes, the Russian State Hydrometeorological University press, St. Petersburg, Russia, 2006. (In Russian).

Kovalenko, V.V.: Using a probability model for steady long-term estimation of modal values of long-term river runoff characteristics, Russ. Meteorol. Hydrol., 39(1), 57–62, doi:10.3103/S1068373914010099, 2014.

Kritsky, S.N. and Menkel, M.F.: On the methods of studying the random variations of river flow, Gidrometeoizdat, Leningrad, USSR, 1946. (In Russian).

Laio, F., Baldassarre, G. D. and Montanari, A.: Model selection techniques for the frequency analysis of hydrological extremes, Water Resour. Res., 45, W07416, doi:10.1029/2007WR006666, 2009.

Lamb, R.: Rainfall-Runoff Modeling for Flood Frequency Estimation, in Encyclopedia of Hydrological Sciences (eds M. G. Anderson and J. J. McDonnell), 11–125 doi:10.1002/0470848944.hsa133, 2006.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M. and Kjeldsen, T.R.: A review of applied methods in Europe for flood-frequency analysis in a changing environment. NERC/Centre for Ecology & Hydrology on behalf of COST, http://nora.nerc.ac.uk/501751/ (last access: 2 November 2017), 2013.

Manton, M.J., Della-Marta, P.M., Haylock, M.R., Hennessy, K.J., Nicholls, N., Chambers, L.E., Collins, D.A., Daw, G., Finet, A., Gunawan, D., Inape, K., Isobe, H., Kestin, T.S., Lefale, P., Leyu, C.H., Lwin, T., Maitrepierre, L., Ouprasitwong, N., Page, C.M., Pahalad, J., Plummer, N., Salinger, M.J., Suppiah, R., Tran, V.L., Trewin, B., Tibig, I. and Yee, D.: Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961–1998, Int. J. Climatol. 21 (3), 269–284, doi:10.1002/joc.610, 2001.

Matalas, N. C. and Wallis, J. R.: Eureka! It fits a Pearson type: 3 distribution, Water Resour. Res., 9(2), 281–289, doi:10.1029/WR009i002p00281, 1973.

Milly, P.C.D., Dunne, K.A. and Vecchia, A.V.: Global pattern of trends in streamflow and water availability in a changing climate, Nature, 438, 347–350, doi:10.1038/nature04312, 2005.

Milly, P., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P. and Stouffer, R. J.: Stationarity is dead: whither water management, Science, 319, 573–574, doi:10.1126/science.1151915, 2008.

Murphy, A.H.: Decision-making models in the cost-loss ratio situation and measures of the value of probability forecasts, Monthly Weather Review, 104, 1058–1065, doi:10.1175/1520-0493(1976)104<1058:DMMITC>2.0.CO;2, 1976.

Murphy, A.H.: The value of climatological, categorical and probabilistic forecasts in the cost-loss situation, Monthly Weather Review, 105, 803–816, doi:10.1175/1520-0493(1977)105<0803:TVOCCA>2.0.CO;2, 1977.

Mylne, K.R.: Decision-making from probability forecasts based on forecast value, Met. Apps., 9, 307–315, doi:10.1017/S1350482702003043, 2002.

Pugachev, V.S., Kazakov, I.E. and Evlanov, L.G.: Basics of statistical theory of automatic system, Mashinostroenie, Moscow, USSR, 1974. (In Russian).

Pearson, K.: Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material. Philosophical Transactions of the Royal Society, 186, 343–414, doi:10.1098/rsta.1895.0010, 1895.

Rosmann, T., Domínguez, E. and Chavarro, J.: Comparing trends in hydrometeorological average and extreme data sets around the world at different time scales, J Hydrol: Regional Studies, 5, 200–212, doi:10.1016/j.ejrh.2015.12.061, 2016.

Rosmann, T. and Domínguez, E.: A Fokker–Planck–Kolmogorov equation-based inverse modelling approach for hydrological systems applied to extreme value analysis, Journal of Hydroinformatics, jh2017079, doi: 10.2166/hydro.2017.079, 2017.

Rogdestvenskiy A.V. and Chebotarev A.I.: Statistical methods in hydrology, Gidrometeizdat, Leningrad, USSR, 1974. (In Russian).

Seibert, J.: Conceptual runoff models - fiction or representation of reality?, comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology, 436, Sweden, 52 pp., 1999.

Sokolovskiy D.L.: River runoff (bases on a theory and methods of calculations). Leningrad, Hydrometeoidat, 540 p. 1968. (in Russian)

Smirnov, N.: Table for estimating the goodness of fit of empirical distributions, Ann. Math. Statist., 19, 279–281, doi:10.1214/aoms/1177730256, 1948.

Shevnina, E.: Deterministic and stochastic models for seasonal forecasting of inflow to reservoirs of hydropower stations, PhD thesis, Russian State Hydrometeorological University, Russia, 188 pp., 2001. (in Russian).

Shevnina, E.: The regional parametrization of the probabilistic model to represent a spring runoff regime in the Russian Arctic, Scientific notes of the Russian State Hydrometeorological University, 21, 38–46, 2011. (in Russian).

Shevnina, E.: Long-term assessment of the multi-year statistical characteristics of the maximal runoff under the climate change over the Russian Arctic, doctor of science thesis, Russian State Hydrometeorological University, Russia, 355 pp., 2015. (in Russian).

Shevnina, E., Kourzeneva, E., Kovalenko, V., and Vihma, T.: Assessment of extreme flood events in a changing climate for a long-term planning of socio-economic infrastructure in the Russian Arctic, Hydrol. Earth Syst. Sci., 21, 2559-2578, doi:10.5194/hess-21-2559-2017, 2017.

Shevnina, E. and Gaidukova E.V.: Hydrological probabilistic model MARCS and its application to simulate the probability density functions of multi-year maximal runoff: the Russian Arctic as a case of study, in: The Interconnected Arctic, Springer International Publishing, 77–87, doi: 10.1007/978-3-319-57532-2_8, 2017.

Taylor, K.E., Stouffer, R.J. and Meehl, G.A.: An overview of CMIP5 and the experiment design, B. Am. Meteorol. Soc., 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

van Gelder, P., Wang, W. and Vrijling, J.K.: Statistical estimation methods for extreme hydrological events, in: New Concepts for Security, NATO Science Series, 78, Springer, Dordrecht, 2006.

Veijalainen, N., Korhonen, J., Vehviläinen, B. and Koivusalo, H.L.: Modelling and statistical analysis of catchment water balance and discharge in Finland in 1951–2099 using transient climate scenarios, Journal of Water and Climate Change, Vol. 3, 55–78, 2012.

Viktorova, N. V. and Gromova, M. N.: Long-term forecasting of characteristics of minimal river runoff discharges in Russia in case of possible climate change, Russ. Meteorol. Hydrol., 33(6), 388–393, 2008.

Vihma, T.: Effects of Arctic sea ice decline on weather and climate: a review, Surv. Geophys., 35, 1175–1214, doi:10.1007/s10712-014-9284-0, 2014.

Volodin, E.M., Dianskii, N.A. and Gusev, A.V.: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, Izvestiya Atmospheric and Oceanic physics, 46(4), 448–466, doi: 10.1134/S000143381004002X, 2010.

Wang, Y. and Zhou, L.: Observed trends in extreme precipitation events in China during 1961–2001 and the associated changes in large-scale circulation, Geophys. Res. Lett., 32 (5), doi: 10.1029/2005GL022574, 2005.

Wagner, A., Lohmann, G. and Prange, M.: Results of modelled Arctic river discharge trends since 7 ka BP, Global Planet Change, 79(1-2), 48–60, doi:10.1016/j.gloplacha.2011.07.006, 2011.

Willmott, C. J. and Robeson S. M.: Climatologically aided interpolation (CAI) of terrestrial air temperature. International Journal of Climatology, 15(2), 221-229, 1995.

WMO-No. 168, guide to hydrological practices. Vol. II: Management of water resources and application of hydrological practices, World Meteorological Organization (WMO), Geneva, Switzerland, 302 pp., 2009.

Yip, Q. K. Y., Burn, D. H., Seglenieks, F., Pietroniro, A. and Soulis, E. D.: Climate impacts on hydrological variables in the Mackenzie River basin, Canadian Water Resource Journal, 37(3), 209–230, 2012.

Table 1. The MARCS model set up: the Iijoki river at Raasakka as a case study.

GRDC ID	River at Gauge	Length, year	m_{1r}, mm year^{-1}	m_{2r}, mm^2 year^{-1}	m_{3r}, mm^3 year^{-1}	$\bar{N}_r$, mm year^{-1}	$\bar{T}_r$ *, °C
6854600	Iijoki at Raasakka (Finland)	100	379	149343	60811610	625	0.2

Notes: m_{1r}, m_{2r}, m_{3r} are the moments of runoff as well as the mean of precipitation ($\bar{N}_r$) were evaluated from observations. The mean air temperature ($\bar{T}_r$)* was not used in the model set up in case of the Iijoki river, however this value allows advancement of the model parametrization (Shevnina et al., 2017).

Table 2. The forcing of the MARCS model for the case study of the Iijoki river at Raasakka.

Global climate model	Climate scenario					
	RCP26		RCP45		RCP85	
	$\bar{T}_{pr}$, °C*	$\bar{N}_{pr}$, mm year^{-1}	$\bar{T}_{pr}$, °C	$\bar{N}_{pr}$, mm year^{-1}	$\bar{T}_{pr}$, °C	$\bar{N}_{pr}$, mm year^{-1}
CaESM2	2.9	673	2.7	652	2.7	652
HadGEM2-ES	1.4	635	2.6	637	2.2	619
INM-CM4	–	–	1.3	645	1.4	660
MPI-ESM-LR	2.5	704	2.2	695	2.9	737

Notes: Projected mean of air temperature ($\bar{T}_{pr}$)* is needed for a regional parametrization scheme (see details Shevnina, 2011), and these values were not used in the model forcing in the case of the Iijoki river at Raasakka. $\bar{N}_{pr}$ is the projected mean of annual precipitation amount.

Table 3. The projected climatology and statistics of annual runoff: a case of the Iijoki river.

Value	Reference period: 1914–2014	Projected period: 2020–2050					
		HadGEM2-ES			MPI-ESM-LR		
		RCP85	RCP45	RCP26	RCP85	RCP45	RCP26
Precipitation, mm $year^{-1}$	625	619	637	635	737	695	704
Specific discharge, mm $year^{-1}$	380	375	386	385	447	421	427
CV	0.19	0.2	0.19	0.19	0.16	0.17	0.17
CS	–0.04	–0.04	–0.04	–0.04	–0.04	–0.04	–0.04
$Q_{10\%}$, m^3s^{-1}	475	473	483	481	527	505	512
$Q_{90\%}$, m^3s^{-1}	293	278	297	296	331	354	359

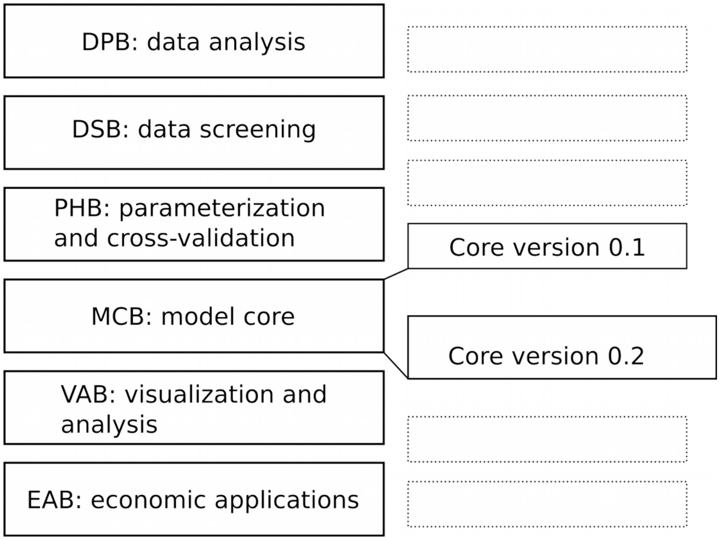

Figure 1: The MARCS^{HYDRO} model structure and core versions.

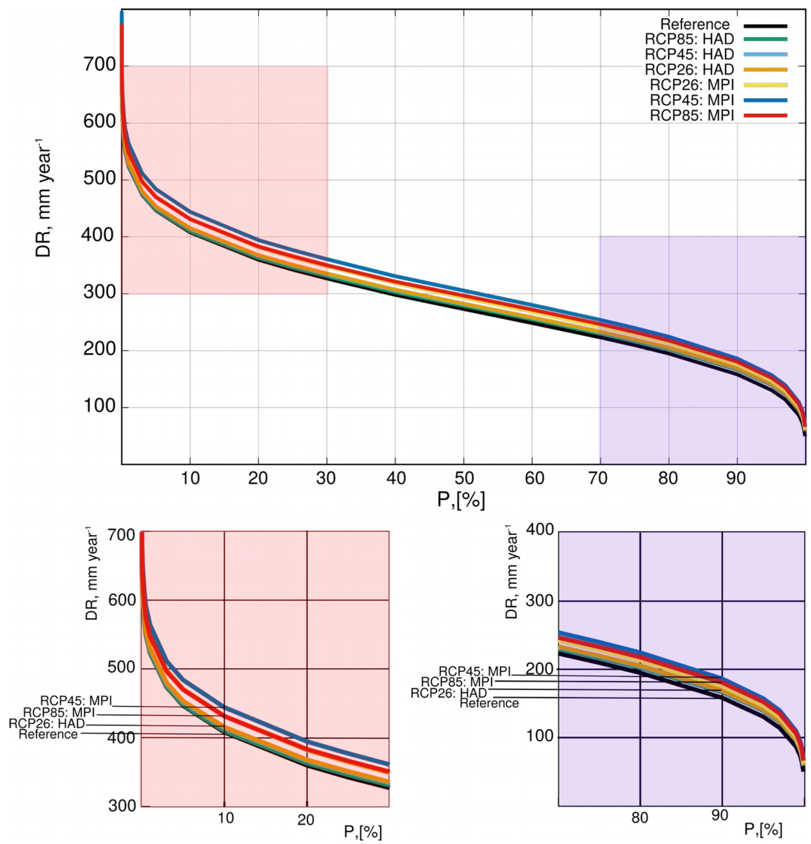

Figure 2: The variability on tails of the EPCs of annual runoff for the reference (black) and projected (colours) periods.