# Large-basin hydrological response to climate model outputs: uncertainty caused by internal atmospheric variability

**A. Gelfan[1,3], V.A. Semenov[2,3,4,6], E. Gusev[1,3], Yu. Motovilov[1,3], O. Nasonova[1,3], I. Krylenko[1,5], and E. Kovalev[1]**

[1] {Water Problems Institute of RAS, Moscow, Russia}

[2] {A.M. Obukhov Institute of Atmospheric Physics of RAS, Moscow, Russia}

[3] {P.P. Shirshov Institute of Oceanology of RAS, Moscow, Russia}

[4]{Institute of Geography of RAS, Moscow, Russia}

[5] {Geographical Department, Moscow State University, Moscow, Russia}

[6] {GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany}

Correspondence to: A. Gelfan (hydrowpi@aqua.laser.ru)

## Abstract

Approach is proposed to assess hydrological simulation uncertainty originating from internal atmospheric variability. The latter is one of three major factors contributing to uncertainty of simulated climate change projections (along with so-called “forcing” and “climate model” uncertainties). Importantly, the role of internal atmospheric variability is the most visible over spatio-temporal scales of water management in large river basins. Internal atmospheric variability is represented by large ensemble simulations (45 members) with ECHAM5 atmospheric general circulation model. Ensemble simulations are performed using identical prescribed lower boundary conditions (observed sea surface temperature, SST, and sea ice concentration, SIC, for 1979-2012) and constant external forcing parameters but different initial conditions of the atmosphere. The ensemble of bias-corrected ECHAM5-outputs and ensemble averaged ECHAM5-output are used as a distributed input for ECOMAG and SWAP hydrological models. The corresponding ensembles of runoff hydrographs are calculated for two large rivers of the Arctic basin: the Lena and the Northern Dvina rivers. A number of runoff statistics including the mean and the standard deviation of annual, monthly and daily runoff, as well as annual runoff trend are assessed. Uncertainties of runoff statistics caused by internal atmospheric variability are estimated. It is found that uncertainty of the mean and the standard deviation of runoff has a significant seasonal dependence on the maximum during the periods of spring-summer snowmelt and summer-autumn rainfall floods. Noticeable non-linearity of the hydrological models' results in the ensemble ECHAM5 output is found most strongly expressed for the Northern Dvina River basin. It is shown that the averaging over ensemble members effectively filters stochastic term related to internal atmospheric variability. Simulated discharge trends are close to normally distributed around the ensemble mean value, which fits well to empirical estimates and, for the Lena River, indicates that a considerable portion of the observed trend can be externally driven.

## 1 Introduction

In river basin hydrology, two groups of approaches are usually applied to assess the impact of changing climate on river runoff. The first group of empirical (data-based) approaches is based on treatment of available hydrometeorological records and includes, for instance, time series analysis of runoff characteristics (see reviews presented by Lins, 2005; Shiklomanov, 2008; Bates et al., 2008), analysis of these characteristics sensitivity to climate variations, particularly by using "elasticity" indices (Sankarasubramanian et al., 2001; Vano and Lettenmaier, 2014), analysis of relationships between spatial and temporal runoff variations ("trading space for time") (Peel and Blöschl, 2011; Singh et al., 2011), etc. The second group includes approaches that are based on hydrological models forced by assigned scenarios of hydrometeorological inputs. These scenarios are constructed either by a transformation of available series of meteorological observations (for example, "delta-change transformation" (Chiew et al., 2009; Motovilov and Gelfan, 2013), "power transformation" (Driessen et al., 2010), or using the global (GCM) and regional (RCM) climate models simulations output (see reviews in Praskievicz and Chang, 2009; Chiew et al., 2009; Peel and Blöschl, 2011; Teutschbein and Seibert, 2010). The latter approach synthesizes up-to-date hydrological models with climate models and provides a better basis to take into account various physical mechanisms of a hydrological system response to the climate change impacts. However, application of this approach is hampered by a number of limitations, first of all, the inconsistency between spatial/temporal resolution of climate models and characteristic scales of hydrological processes in river basin, which differ by several orders of magnitude, both in time and space (Blöschl and Sivapalan, 1995)). Another serious limitation is related to climate models' capability to accurately reproduce variability and the mean state for many meteorological characteristics, especially for precipitation (see, for example, Kundzewicz et al., 2008; Kundzewicz and Stakhiv, 2010; Anagnostopoulos et al., 2010). Explosive increase in computing resources occurred during the last years, development of measuring technologies and methods of data processing, as well as numerical methods, favor for the improvement of climate models, significant increase in their productivity and spatial resolution, and better simulation of regional climate (Flato et al., 2013). This promotes a wider usage of the model-based approach for assessment of the climate change impact on river runoff. However, a significant uncertainty in these assessments still remains and their interpretation should be considered with caution, especially for practical applications in the field of long-term water management (Wilby, 2010; Kundzewicz and Stakhiv, 2010).

Part of the total uncertainty inherent to assessments of climate change hydrological consequences is caused by limitations of our knowledge about the dynamics of climatic and hydrological systems, nature of their interrelationships, insufficiency of measured data, etc., and, potentially, can be reduced with increasing understanding of these systems (epistemic uncertainty). Another part of this uncertainty is a structural one, which does not depend on acquiring new knowledge and data and is an inherent property of these systems. Evaluation of this structural, inherent uncertainty impact is the key issue to realize the potential to obtain reliable assessments of climate-driven changes in river runoff (see, e.g., discussion in Koutsoyiannis et al., 2009).

Uncertainty of assessments of hydrological response to climate change is primarily caused by uncertainty of the future climate projections. The latter is related to three independent factors (Hawkins and Sutton, 2009; Deser et al., 2012). The first, so-called "response uncertainty" or "model uncertainty", is caused by differences in climate response to identical external (e.g., anthropogenic) forcing in different climate models. The model uncertainty arises from structural differences (in particular spatial resolution) between climate models, different parameterizations of physical processes, numerical methods, etc., related to scientific advances in understanding and description of a climate system and therefore can be potentially reduced. The second factor is so-called "scenario uncertainty" and represents uncertainties related to prescribed scenarios of future anthropogenic greenhouse and aerosols emissions. The third factor is the internal, natural variability of climate system (or so-called "climatic noise"), which exists also in the absence of external forcing and results from stochastic nature of atmospheric dynamics, its instability to small perturbations, and also internal (often non-linear) modes of variability in the atmosphere and the ocean at different time and spatial scales. Climatic noise is a major source of physically based structural uncertainty in climate change projections and it determines a lower limit of uncertainty that can be reached in climate system modeling (Braun et al., 2012).

The major components of climatic noise are stochastic fluctuations in the atmosphere and the ocean. Large heat capacity, relatively low ocean circulation velocities (relative to atmosphere) and existence of internal oscillatory modes with (quasi) periodicity ranging from years to centuries (Semenov et al., 2010; Latif and Keenlyside 2011, Latif et al., 2013) provide a certain predictability of oceanic processes. This so-called “second kind of predictability”, particularly predictability on time scale of about ten years that has been recently found to be potentially approached by modern climate models, is currently an object of intense research (e.g., Latif and Keenlyside 2011). Another source of uncertainty is caused by internal atmospheric variability and related to stochastic dynamics of atmosphere, instability of atmospheric circulation to small perturbation of parameters. Commonly known as the “butterfly effect”, this kind of instability was illustrated in the classical work by Edward Lorenz (1963). Such an uncertainty determines a time limit for a weather forecast that does not exceed two weeks and leads to essentially different realizations of the atmospheric state beyond this limit given the same boundary and external forcing but small (within the measurement error) changes in initial conditions. Hereinafter, we use the term “climatic noise” to refer only to this kind of uncertainty caused by internal atmospheric variability. Our study focuses on transformation of the climatic noise by hydrological models and its impact on the uncertainty of simulated runoff. Note that the role of the climatic noise is most important on time scales from years to first decades and on regional spatial scales (Räisänen, 2001; Hawkins and Sutton, 2009), i.e. on the spatial-temporal scales of water resource management in large river basins.

Analysis of uncertainty related to internal atmospheric variability is based on ensemble climate model simulations with identical external forcing and different initial conditions (“multireplicate ensemble”). This approach results in ensemble of realizations or trajectories of climate system states that differ from each other solely due to internal variability (Yip et al., 2011; Braun et al., 2012; Deser et al., 2012; Sansom et al., 2013; Semenov, 2014). To obtain reliable statistical assessments of variability within an ensemble, it is necessary to calculate several dozens of simulation trajectories as a minimum. Such calculations using GCM require large computational resources. Simulations with climate models participating in the World Climate Research Programme (WCRP) Coupled Model Intercomparison Project Phase 3 and 5 (CMIP3 and CMIP5) (Meehl et al., 2007; Taylor et al., 2012) used for the 4th and 5th IPCC assessment reports respectively include just a few (usually not exceeding ten) trajectories for any particular model (Peel et al., 2014). This fact is partially responsible for the absence, till recently, of studies of climate noise effect on assessments of uncertainty in river runoff climate-driven changes. The first publications in this field appeared, to our knowledge, in 2014 (Seiller, Anctil, 2014; Lafaysse et al., 2014; Peel et al., 2014).

Seiller and Anctil (2014) constructed climate scenarios using Canadian GCM (CGCM) with spatial resolution of 3°×3.75° followed by dynamic downscaling of the calculated data to a local scale with resolution of 45 km. Ensemble of realizations calculated under different initial conditions for simulating climate system internal variability consisted of 5 members. The realizations were assigned as an input for 20 conceptual runoff models with lumped parameters to calculate river runoff in a small, around 30 km2, basin in the south-west of Canada. The authors demonstrated that the uncertainty of river runoff assessments caused by climate noise exceeds the uncertainty of hydrological models.

To increase the climate scenarios ensemble size, which simulates internal variability, Lafaysse et al. (2014) used stochastic generators and assigned the constructed stochastic scenarios as an input into ISBA/Durance land surface model. Similar approach was presented by Peel et al. (2014) to increase the number of climatic trajectories simulated by five GCMs. The authors developed a stochastic procedure to generate time series of monthly meteorological variables with statistics close to those obtained from GCM simulations. The generated hundred of 250-year meteorological time series were used to force the conceptual PERM hydrological model.

On the one hand, the use of stochastic generators for calculating a large ensemble of climate system trajectories is a much more efficient (from the computational point of view) approach to assess climate-driven changes in river runoff when compared to simulation of GCM-realizations ensemble (Hawkins and Sutton, 2009; Yip et al., 2011; Deser et al., 2012; Sansom et al., 2013). On the other hand, the applied stochastic procedures create an additional and ambiguously interpreted source of uncertainty.

In this paper we have tried to assess, using physically based hydrological models, the uncertainty in simulated river runoff characteristics of large river basins taking into consideration internal variability of the atmosphere. The latter was simulated in a large (45 members) ensemble of GCM-realizations of the current climate period (1979-2012) initialized under different initial conditions but using identical boundary forcing (sea surface temperatures and sea ice concentrations). Case studies were carried out for two large watersheds of the Arctic basin: the Lena River (catchment area F=2 488 000 km$^2$) and the Northen Dvina River (F= 357 000 km$^2$). We emphasize that our study focuses on present day climate variations with relatively smaller contribution of the external forcing compared to studies considering future climate projections to the end of the 21$^{st}$ century (e.g., Déqué et al., 2007; Hagemann et al., 2009). On such time scales, the impact of internal variability diminishes compared to other uncertainty sources.

The paper is structured as follows. Section 2 presents the main physiographic and climatic characteristics of the basins under consideration. Further a short description of the used hydrological models ECOMAG and SWAP can be found, as well as the results of their validation against hydrological observations in the basins under study. Section 4 contains a brief description of the atmospheric general circulation model (AGCM) ECHAM5, the design and results of numerical experiments on simulating internal atmospheric variability. In Section 5, runoff characteristics uncertainty caused by internal atmospheric variability is analyzed on the basis of the simulated runoff ensemble. Uncertainties of the mean and the variance of the river discharge averaged over different time intervals (calendar day, calendar month, year), as well as to the uncertainty in long-term trend of the simulated annual discharge are emphasized. The last section summarizes the results and presents the main conclusions.

## 2 Study basins and datasets

The case studies were carried out for two Arctic river basins: the Lena River and the Northern Dvina basins. The Lena River is one of the largest rivers in the Arctic that flows northward from mid latitudes to the Arctic Ocean (Fig. 1), and it contributes about 15% of total freshwater flow into the ocean. The basin occupies an area of 2 460 000 km$^2$ extending from 103°E to 142°E and from 52°N to 74°N. The length of the basin from the South to the North is more than 2400 km; its average width is about 2000 km. There are four main types of landscapes within the Lena River basin: Arctic wilderness, tundra, forest tundra and taiga forests, which occupy almost 70% of the basin area. The main part of the basin has mountain relief with heights ranging in general from 600 to 2000 m (reaching 3500 m in the southern part of the basin). The climate is extremely continental, with surface air temperatures being extremely low in winter (as cold as -50 – -65 $^{o}$C) and high in summer (up to +20 -+35 $^{o}$C). The whole territory of the basin is located in the permafrost zone. The Lena River runoff is characterized by spring-summer snowmelt flood, summer and autumn rain floods and extremely low water levels in winter. Maximum discharge of 189 000 $m^3$/s was observed at the basin outlet Stolb – on June the 1st, 1984. The average annual discharge of the Lena River is about 15 370 $m^3$/s. There are over 80 meteorological and over 20 runoff hydrological stations within the basin.

The Northern Dvina River basin with an area of 360 000 $km^2$ occupies vast flat forested territory in the northern part of East European plain from 39°E to 56°E and from 58°N to 66°N and flows northward to the White Sea basin. Taiga forest covers more than 80% of the river basin with the northern part changing for tundra landscapes. The climate of the territory is influenced by cyclonic activity. Precipitation exceeds evaporation which leads to excessive wetness. More than 60% of the annual runoff belongs to spring flood period. Maximum discharge of 36 200 $m^3$/s was observed at the basin outlet Ust-Pinega on 28th of April 1953. The average annual discharge of the Northern Dvina River is about 3400 $m^3$/s. There are 35 meteorological and over 10 runoff hydrological stations within the basin.

Due to low anthropogenic burden and absence of reservoirs for regulating the main river flow, the Northern Dvina and the Lena River basins are good objects for case studies aimed to estimate runoff response to climate variations.

## 3 Hydrological models

Two hydrological models, ECOMAG (Motovilov et al., 1999a) and SWAP (Gusev and Nasonova, 1998), developed at the Water Problems Institute of RAS (Moscow) are used in this study. These models have been successfully tested against observation data all over the world.

Physically-based semi-distributed model ECOMAG (ECOlogical Model for Applied Geophysics) developed by Yu. Motovilov (Motovilov et al., 1999a) was earlier applied for hydrological simulations in many river basins of various sizes and located in different natural conditions: from small-to-middle size Scandinavian basins (e.g. Motovilov et al., 1999b) to the great Volga and Lena Rivers with watershed areas exceeding a million km$^2$ (Gelfan and Motovilov, 2009; Motovilov and Gelfan, 2013). Since 2004, the ECOMAG model has been utilized in an operational mode for hydrological characteristics and water inflow simulation into the Volga-Kama and the Angara-Yenisey reservoir cascades in Russia, which are among the largest reservoir cascades worldwide.

Physically based land surface model Soil Water-Atmosphere-Plants (SWAP) developed by Ye. Gusev and O. Nasonova (Gusev and Nasonova, 1998) was intensively validated, in particular, within several model intercomparison projects (PILPS, Rhone-AGG, MOPEX, SnowMIP, GSWP-2) for different river basins and experimental sites located in various natural zones (from areas in tropical zone to regions with permafrost) and characterized by different spatial scales (from small experimental sites and catchments to the whole land surface of the Earth). The results of the model testing are presented, particularly, in (Gusev and Nasonova, 1998, 2003; Gusev et. al, 2011)

Both models describe interception of rainfall/snowfall by the canopy, processes of snow accumulation and melt, soil freezing and thawing, water infiltration into unfrozen and frozen soil, evapotranspiration, thermal and water regime of soil, overland, subsurface and channel flow. ECOMAG model utilizes semi-distributed approach with the whole river basin interpreted as a number of sub-basins. It takes into consideration topography, soil and land cover characteristics of a particular sub-basin. For each sub-basin, hydraulic properties of soil as well as land-cover properties are scaled taking into account sub-basin area (Motovilov et al. 1999a, b). Subsurface and groundwater routing is based on the Darcy law, while the surface runoff and channel flow are described by a kinematic wave equation. SWAP model utilizes a regular spatial grid with a size of $1^{o}\times1^{o}$. The cells are connected into channel network. Streamflow transformation within the network is calculated with the use of a linear model using TRIP algorithm (Oki et al., 1999)

Most of the parameters are physically meaningful and can be assigned from literature or derived through available measured characteristics of topography, soil, and land-cover. Some key-parameters of the models are calibrated against streamflow measurements and, if available, measurements of the internal basin variables (snow characteristics, soil moisture, groundwater level, etc.).

ECOMAG model is forced by daily time series of air temperature, air humidity and precipitation. The SWAP inputs include 3-hour data of incoming radiation, precipitation, air temperature and humidity, atmospheric pressure, and wind speed. The forcing data can be taken from meteorological observations or GCM-outputs.

Both models were applied earlier for simulating runoff hydrographs based on multi-year hydrometeorological observations in the Lena and the Northern Dvina River basins and demonstrated good performance of simulations (Motovilov and Gelfan 2013; Gusev et. al, 2011; ; Krylenko et al., 2014, Gusev et al., 2015). Trial-and-error manual procedure and Shuffle Complex Evolution (SCE-UA) automatic algorithm were applied for calibration of ECOMAG and SWAP, respectively. Widely-used Split-Sample Test (Klemeš, 1986) was utilized for model validation. Both calibration and validation procedures were carried out against daily streamflow data measured at several gauges of these large basins. Nash and Sutcliffe (1970) efficiency, NSE, and bias evaluation criteria were adopted to summarize the goodness of fit of simulated and measured daily discharge series. As an example, the evaluation criteria calculated for the outlets of the Lena and the Northern Dvina River basins and adopted from (Motovilov and Gelfan 2013; Gusev et. al, 2011; ; Krylenko et al., 2014, Gusev et al., 2015) are shown in Table 1.

## 4 Atmospheric general circulation model description and inernal variability simulations

Ensemble simulations were performed with atmospheric general circulation model (AGCM) ECHAM5 developed at the Max Planck Institute for Meteorology (Roeckner et al., 2003). This model is a climatic version of AGCM based on spectral weather forecast model of the European Centre for Medium-Range Weather Forecasts (ECMWF) that employs state-of-the-art physics. The model version used here has a horizontal resolution of T63 (1.8° × 1.8° latitude × longitude) and 31 vertical levels. All 45 ensemble simulations use identical prescribed lower boundary conditions at atmosphere-ocean interface. These conditions are taken from HadISST1.1 (Hadley Centre, UK) dataset that consists of global empirical analysis of the sea surface temperature (SST) and the sea ice concentrations (SIC, a portion of model grid cell covered by sea ice) (Rayner et al. 2003). The simulation period is from 1979 to 2012. The start of simulations in 1979 was motivated by beginning of the era of continuous satellite monitoring of the sea ice cover that provides most reliable SIC data. This is important for correct simulations of the climate in high-latitudes (Semenov and Latif, 2012). Greenhouse gas concentrations in the model are kept constant and represent modern climate conditions

(348 ppm for $CO_2$, and 1.64 ppm for methane). All other external forcing parameters (such as orbital parameters, solar radiation, other radiative-active gases and aerosols) also correspond to modern climate conditions and do not vary. The only differences between the simulations are initial conditions of the atmosphere (model atmospheric state on the 1st of January, 1979) that are prescribed as instant atmospheric states at different 12 hour intervals in December 1978. Thus, the ensemble consists of 45 simulations with identical boundary and external forcing but different initial conditions. Note that the characteristics at atmospheric lower boundary over land (soil temperature and moisture, snow cover) are computed by AGCM using a land surface model and simulated heat and water fluxes (Roeckner et al. 2003).

Such ensemble simulations with time-varying SST and SIC according to observational data allow one to estimate a contribution of the varying SST and SIC fields to the observed changes in atmospheric characteristics (the mean, trends, variability) during the simulation period (assuming that AGCM correctly reproduces a response to varying boundary conditions). When considering changes of atmospheric variables consisting of changes caused by external to atmosphere factors (SST and SIC) that are supposed to be the same in all simulations and internal variability (due to stochastic atmosphere dynamics and thus independently distributed), the averaging over large ensemble members effectively filters stochastic terms (climatic noise) and results in an estimate of the external signal related to SST and SIC changes. Similar approach will be applied in section 5.3 to estimate externally forced part of long-term changes in hydrological characteristics that provides a basis for estimating potential predictability limits for hydrological systems.

To illustrate differences between individual ensemble members arising from internal atmospheric dynamics, several meteorological characteristics were averaged over the Lena River catchment area. Figs. 2 (top) show ensemble (45 realizations) of the mean annual temperature and precipitation for the period of simulations (1979-2012); Figs. 2 (bottom) demonstrate ensemble of the mean daily values of these variables averaged over the simulation period.

A positive trend for both temperature and precipitation (Fig. 2 top) agrees with global warming and the tendency of precipitation increase in high northern latitudes accompanying temperature increase. Intra-ensemble standard deviations of the annual temperature and precipitation values caused by internal stochastic atmospheric dynamics account for 0.5°C and 0.08 mm/day respectively. The standard deviations of the daily mean temperature vary within 0.4-0.8 °C during a year, while the deviations of precipitation are about 0.02-0.04 mm/day in winter and reach as much as 0.30 mm/day for some summer days. The following section will address a question how such uncertainty is transformed to uncertainty in river discharge.

An important factor that should be taken into account while analyzing ECHAM5 simulations is a model bias (e.g., Hagemann et al., 2006; 2013). Even when forced with observed fields of SST and SIC, ECHAM5 simulates the mean climate over land areas that differs from observations of the corresponding period. The sources for model bias include deficiencies in parameterizations and incomplete description of some physical processes, numerical schemes, low model resolution (Flato et al., 2013, IPCC AR5). In our experiments, ECHAM5 simulates a colder winter climate in high latitudes of the Northern Hemisphere that is related to higher sea level pressure over the Arctic and weakened zonal flow in mid and high latitudes (not shown).

A post-processing procedure, analogous to that proposed by Velázquez et al. (2013), was applied to correct biases in ECHAM5-outputs before using them as inputs into hydrological models. The correction factors were computed based on the difference between the ensemble-mean climate variables modelled for the reference period (1979-2009) and corresponding observed variables averaged over the basin areas under consideration. The correction factors were then added to ECHAM5-simulated 6-hour meteorological fields. Comparison of the spatial fields of mean annual values of precipitation, air temperature and humidity obtained from data registered in the meteorological stations located within the Lena River basin and processed from the simulated data is illustrated, as an example, by Fig. 3. Figure 3 shows that the applied post-processing allowed us to obtain rather similar patterns of the above listed variables taking into account sparseness of the meteorological monitoring network in the basin. In addition, the similarity is rather surprising taking into account that the assigned correction factor is based on the model-observation differences averaged over very large basin. Thus, ECHAM5 demonstrates good performance in simulating spatial distribution of deviations from the basin averaged values of precipitation, air temperature and humidity.

## 5. Experiment design, results and discussion

Due to stochastic nature of climate, hydrological models cannot provide predictions of specific streamflow hydrograph series (even for the past, not to mention for the future) on the basis of the climate model outputs. In other words, hydrological models operating on outputs from climate model are confined, similar to climate models, to making projections rather than predictions (Refsgaard et al, 2014) and are able to provide only information on statistical characteristics of runoff series. Below we present approaches to and results of estimating these statistical characteristics from simulated ensembles of multi-year streamflow hydrographs, as well as analysing uncertainty of the estimations.

An ensemble of NI=45 time series of meteorological variables simulated by ECHAM5 for the period of NY=34 years (from 1.01.1979 to 31.12.2012) was assigned as a distributed input into ECOMAG and SWAP hydrological models. With the help of these two models, 45-member ensembles of daily streamflow series each of 34-year length were calculated for the Lena River and the Northern Dvina River. From these hydrograph ensembles, the mean values and the standard deviations of annual, monthly and daily runoff were estimated. Then, 95% confidence intervals for the estimates were calculated as an indication of uncertainty in these estimates caused by the internal variability of the atmosphere. Whilst calculating the confidence intervals, it was assumed that these estimates followed the Gaussian probability distribution.

More precisely, the estimates were calculated as follows. Assume a calculated water discharge be $X_{ij}$, where $i$=1,2,…,45 is the realization number referred to the assigned initial conditions in the climate model; $j$=1,2,…34 is the number of year within the simulation period. In this study, $X_{ij}$ can be either an annual discharge for a specific year, or a monthly discharge for a specific calendar month, or a daily discharge for a specific calendar day, derived from $i$-th realization and related to $j$-th year. Thus, according to the experimental design, any variable, be it annual, monthly or daily, is considered as 45×34 matrix (for instance, matrix of January discharges or matrix of July 25 discharges).

To obtain the above mentioned statistical characteristics and their confidence intervals, the following formulae were used:

$M$ -estimate of the mean value:

$$M = \frac{1}{(NY \times NI)} \sum_{i=1}^{NY} \sum_{j=1}^{NI} X_{ij} \qquad (1)$$

$SD$ -estimate of the standard deviation:

$$SD = \sqrt{\frac{\sum_{i=1}^{NY}\sum_{j=1}^{NI}\left(X_{ij} - M\right)^2}{\left(NY \times NI\right) - 1}} \qquad (2)$$

the confidence interval $\gamma_M$ for $M$ :

$$\gamma_M = \left( M + \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\sigma_M ; M - \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\sigma_M \right) \qquad (3)$$

the confidence interval $\gamma_{SD}$ for $SD$ :

$$\gamma_{SD} = \left( SD + \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\sigma_{SD} ; SD - \Phi^{-1}\left(\frac{1+\alpha}{2}\right)\sigma_{SD} \right) \qquad (4)$$

where $\alpha$ is the confidence probability, $\Phi^{-1}(x)$ is the inverse of the cumulative normal distribution function; $\sigma_M$ is the standard deviation of $M$ , equal to

$$\sigma_M = \sqrt{\frac{\sum_{i=1}^{NI}\left(M_i - M\right)^2}{NI - 1}}, \qquad (5)$$

$M_i = \frac{1}{NY}\sum_{j=1}^{NY} X_{ij}$ ;

$\sigma_{SD}$ is the standard deviation of $SD$ , equal to

$$\sigma_{SD} = \sqrt{\frac{\sum_{i=1}^{NI}\left(SD_i - M_{SD}\right)^2}{NI - 1}}, \qquad (6)$$

$SD_i = \sqrt{\frac{1}{(NY-1)}\sum_{j=1}^{NY}\left(X_{ij} - M_i\right)^2}$ , $M_{SD} = \frac{1}{NI}\sum_{i=1}^{NI} SD_i$

Hereafter, the confidence intervals of estimates are evaluated for $\alpha$ =95% confidence probability, i.e. $\Phi^{-1}\left(\frac{1+0.95}{2}\right)$=1.96

To compare uncertainty in statistical estimates of runoff characteristics, which differ in their absolute value, normalized widths of the confidence intervals were used. Uncertainty indices $UN(M)$ and $UN(SD)$ of $M$ and $SD$ estimates, respectively, are introduced which are considered to be half of the width of 95% confidence interval of the corresponding estimate divided by its mean value, i.e.:

$$UN(M) = \frac{\gamma_M^+ - \gamma_M^-}{2M} = \frac{1.96\sigma_M}{M} \tag{7}$$

$$UN(SD) = \frac{\gamma_{SD}^+ - \gamma_{SD}^-}{2SD} = \frac{1.96\sigma_{SD}}{SD} \tag{8}$$

where $\gamma_\bullet^+$ and $\gamma_\bullet^-$ are the right and the left limits of the confidence interval, respectively.

In the next Subsections, the results of ensemble simulation of river runoff for each study basin are presented. First, analysis of the uncertainty indices of annual, monthly and daily mean estimates for both river basins is presented. The calculated estimates are compared to the corresponding observation data estimates. Then, analogous results are shown for standard deviations of runoff values averaged over the same time intervals (a year, a month, a day). Finally, uncertainty in trends in annual runoff is calculated and discussed.

## 5.1 Estimates of the mean runoff and their uncertainty

Table 2 demonstrates the uncertainty indices $UN(M)$ for $M$-estimates of annual, monthly and daily runoff calculated by formula (7). Intra-annual variation of uncertainty indices $UN(M)$ for $M$-estimates of daily runoff is shown in Fig. 4.

The following conclusions can be derived based on the presented results:

(1) Uncertainty in the mean runoff values calculated by both of the models for both rivers decreases with the increasing interval of runoff averaging.

It is shown in Table 2 that the uncertainty index $UN(M)$ for the $M$-estimates of daily runoff varies from 8% to 24% ; $UN(M)$ for monthly runoff - from 7% to 19%; $UN(M)$ for annual runoff – from 6% to 10% depending on the model used and the study river basin. However, uncertainty indices for monthly and daily runoff estimates have distinguished seasonal variations, and maximum values of the uncertainty considerably exceed their average values within a year. For example, as can be seen from Fig. 4, uncertainty index $UN(M)$ for daily runoff can be more than 50%, and $UN(M)$ for mean monthly runoff estimates reaches 41% (Table 2).

(2) Internal atmospheric variability has maximal influence on uncertainty in the estimates of the mean runoff during snowmelt/rainfall flood periods for both rivers. Uncertainty of estimates of the mean runoff during winter months is small.

Uncertainty indices $UN(M)$ for $M$-estimates of monthly runoff during the period of snowmelt floods and rainfall floods amount to 21-24% for the Lena River and 35-41% for the Northern Dvina River depending on the applied hydrological model (see Table 2). The uncertainty $UN(M)$ for daily runoff is even greater (Fig. 4): for snowmelt flood this value is 42-55% for both rivers. Uncertainty $UN(M)$ for monthly runoff during winter periods is much less (2-13% for the Lena River and 2-19% for the Northern Dvina River); the same applies to daily runoff during winter (see Fig. 4). Possible explanation of these findings is that physical mechanisms of flood events are more sensitive to intra-ensemble changes of the climate model outputs than more inertial mechanisms of low flow generation.

(3) Uncertainty in the mean runoff estimates for the Lena River basin appeared to be significantly less than the ones for the Northern Dvina River when using both models. Moreover, intra-annual irregularity of $UN(M)$ is more noticeable for the Northern Dvina simulations both on monthly (Table 2) and daily (Fig. 4) time scales. In other words, the Northern Dvina simulated hydrographs appeared to be more sensitive to the atmospheric variability.

This difference of uncertainty in the mean runoff estimates is related to peculiarities of river runoff generation in the study basins. These peculiarities can manifest themselves, for example, in a degree of non-linearity of river basin response to climate impact: increase in non-linearity, generally speaking, should lead to increase in the uncertainty in the calculated runoff characteristics. Therefore, one can assume that the mechanisms of runoff generation and transformation of climate impact on variations of river runoff are more linear in the Lena River basin than in the Northern Dvina River basin. To validate this assumption, we compared two mean hydrographs for each basin. One was calculated by averaging over the ensemble of 45 simulated mean hydrographs (an averaged response to ensemble input) and the other simulated by the hydrological models using one meteorological input obtained by averaging over 45 ECHAM5-outputs (a response to the ensemble averaged input).

If the response of the hydrological system to climate impact is linear, these hydrographs should be similar, whereas non-linearity should lead to an increased difference between these hydrographs. The results of the comparison are shown in Fig.5.

As one can see from Fig. 5, both models show that the response of the hydrological system of the Lena River basin is close to linear, while the response of the Northern Dvina River is essentially non-linear. This supports the above mentioned assumption about an increased effect of internal atmospheric variability on uncertainty of the mean river runoff estimates in the Northern Dvina River basin due to a greater non-linearity of the mechanisms of runoff generation compared with the Lena River basin. Note, that due to the effect of averaging, peak discharge of the ensemble mean hydrographs is always lower than the hydrograph peak simulated from the mean outputs (see Fig. 5).

(4) Uncertainty of the mean runoff estimates determined using different models vary insignificantly, despite the fact that these models require different input data, are differently structured and parametrized. Thus, the average uncertainty indices $UN(M)$ for SWAP-simulated monthly runoff are 11% for the Lena River and 19% for the Northern Dvina River; when using ECOMAG, the values are 7% and 19%, respectively.

As the next step, we compare the obtained $M$-estimates of the simulated runoff with the corresponding estimates derived from streamflow observations in the basins under consideration.

Figs. 6 and 7 present a comparison between $M$-estimates for annual, monthly and daily discharges calculated at the basin outlets with the corresponding estimates obtained from the time series of the discharges observed for 31 years (1979-2009). These Figures also show 95% confidence intervals $\gamma_M$ for the calculated estimates of the mean values computed by formulae (3) and (5).

The comparison of the calculated estimates with the mean runoff characteristics evaluated by available observational series has demonstrated that calculation errors, when using both models, increase with decreasing time interval of discharge averaging. Estimates of the mean annual runoff are characterized by the smallest error: 5% and 18% for the Lena River, and 10% and 33% for the Northern Dvina River depending on the hydrological model used. The errors of the mean monthly and the mean daily runoff estimates are usually much greater. It is especially noticeable for the periods of spring-summer snowmelt flood and summer-autumn rainfall floods for both rivers: error of the mean monthly runoff can reach several dozens of percent, and for the mean daily runoff – hundreds of percent. Winter months are an exception with errors for both mean monthly and mean daily runoff usually not exceeding 30-40%. It turned out that all calculated estimates of mean runoff were closer to the corresponding estimates based on empirical data for the Lena River than for the Northern Dvina River. This can be explained by a weaker natural variability of the runoff characteristics at a larger basin of the Lena River.

### 5.2 Estimates of the standard deviation of runoff and their uncertainty

While analyzing $SD$-estimates of runoff, we focused on the same issues, which were discussed in the previous sections when analyzing the corresponding $M$-estimates. Specifically, we considered dependence of uncertainty indices $UN(SD)$ on the interval of runoff averaging, intra-annual changes in $UN(SD)$, difference in $UN(SD)$ for different basins, and comparison of the $SD$-estimates with the corresponding estimates calculated from the available observed streamflow time series.

Table 3 presents the uncertainty indices $UN(SD)$ for $SD$-estimates of annual, monthly and daily runoff at the outlet of the studied rivers, which were calculated by equation (8). Intra-annual variation of the uncertainty indices $UN(SD)$ for daily runoff $SD$-estimates is shown in Fig. 8.

A comparison of the uncertainty indices estimates for the standard deviation (Table 3, Fig. 8) and the mean (Table 2, Fig. 4) reveals that the uncertainty indices $UN(SD)$ for $SD$-estimates of runoff characteristics are, unsurprisingly, much higher than the uncertainty $UN(M)$ for $M$-estimates for the same runoff averaging interval. Similar to $UN(M)$, an uncertainty trend of the standard deviation can be noticed when the time averaging interval of water discharge decreases: $UN(SD)$ increases from 24-31% for annual runoff to 30-52%, as the average, for monthly runoff, and 36-98% for daily runoff.

At the same time uncertainty in $SD$-estimates of monthly and daily water discharges significantly varies within a year, and the maximum values of the uncertainty index $UN(SD)$ for these estimates considerably exceed their mean values. For example, $UN(SD)$ for some calendar months is close to 100% (Table 3), and $UN(SD)$ for daily runoff estimates reaches hundreds of percent (Fig. 8).

Similar to the results for the uncertainty of the mean runoff estimates, the impact of atmospheric variability on standard deviation uncertainty has a significant intra-annual variation. Uncertainty $UN(SD)$ for monthly and daily runoff reaches its maximum in the periods of spring-summer snowmelt floods and summer-autumn rainfall floods at both rivers (see Table 3 and Fig. 8). Uncertainty $UN(SD)$ for winter runoff is somewhat smaller but still large in contrast to the uncertainty in the mean values during winter months, which, as it was shown above, significantly decreases. This result can be explained by a small variation of winter runoff.

Uncertainty indices $UN(SD)$ for $SD$-estimates of the Lena River runoff for both hydrological models are smaller than for the Northern Dvina River (which is similar to results for $UN(M)$). Uncertainty in annual runoff varies very slightly (24-36% for the Lena River and 30-31% for the Northern Dvina River). However, the decrease of the averaging interval to a month and a day leads to a significant increase in $UN(SD)$ variations for both basins. As it was shown above, the difference of $UN(SD)$ values can be accounted for stronger non-linearity of the runoff generation mechanisms for the Northern Dvina River than for the Lena River.

Figs. 9 and 10 show comparison of $SD$-estimates for annual, monthly and daily discharges calculated at the basin outlets with the corresponding estimates obtained from the observed time series of the discharge for the period 1979-2009. These Figures also present 95% confidence intervals $\gamma_{SD}$ of the standard deviation calculated estimates (according to equation (4)).

The calculations showed that the relative errors of the $SD$-estimates derived by simulated runoff time series were fairly large in comparison with the corresponding estimates based on empirical data. These estimates were most similar for the annual runoff: 3% and 21% for the Lena River and 41% and 57% for the Northern Dvina River depending on hydrological model. When the time averaging interval for water discharge decreases, errors in the estimates increase for both models and both rivers, which is particularly noticeable for the winter season, when the $SD$-estimates are sometimes hundreds percent lower in comparison with their observed variability. It should be noted that large relative errors may result from the small absolute differences due to very small discharge values in winter season. Similar to $M$-estimates, $SD$-estimates are closer to corresponding estimates based on empirical data for the Lena River than for the Northern Dvina River.

### 5.3 Estimate of annual runoff trend and its uncertainty

As it has been already discussed in Section 4, averaging over the ensemble of simulated realizations allowed us to filter off a random component caused by atmospheric variability and to assess the impact of the “signal” caused by factors external to atmosphere (related to the prescribed observed SST and SIC changes in our experiments). Such an assessment is presented in this Subsection with an analysis of long-term annual runoff trend.

Fig. 11 shows long-term variations of the annual discharge values observed at the outlets of both rivers compared with the corresponding values averaged over the ensemble of 45 runoff hydrographs realizations calculated using ECOMAG model.

It is shown that individual realizations of calculated annual discharges differ from each other and are, in general, only slightly correlated with corresponding observed time series. For the Lena River simulations, correlation coefficients vary from -0.31 to of 0.56 with the mean value of 0.17. Note that correlation between the observed annual discharges and the ensemble mean annual discharges is rather high (0.51). However, the standard deviation of the observed discharge time series (17 616 $m^3/s$) is almost 1.3 orders greater than that of the mean ensemble discharge time series (901 $m^3/s$). It is necessary to mention, that corresponding correlations derived from SWAP simulation experiments are very close to ones listed above: correlation coefficients vary from the minimum of -0.29 to the maximum of 0.53 with the mean value of 0.14.

For the Northern Dvina River, correlation coefficients between individual realizations and the observed annual discharge series are, mostly, statistically insignificant under a reasonable significance level. The coefficients vary from the minimum of -0.56 to the maximum of 0.30 with the mean value of -0.04. The correlation coefficient between the observed annual discharges and mean ensemble annual discharges is also insignificant (-0.19). Again, corresponding correlations derived from the SWAP simulation experiments are very close to those obtained by ECOMAG simulations: correlation coefficients vary from the minimum of -0.40 to the maximum of 0.33 with the mean value of -0.03, as well as correlation between the observed annual discharges and the mean ensemble annual discharges calculated by SWAP model is insignificant and equals to -0.13.

Fig. 12 shows histograms of linear trends of annual runoff obtained for each realization from the calculated ensembles. The trend calculated from the observational data (Slope(fact)) and the mean trend calculated by averaging over 45 trends for the individual realizations (Slope(mean calc)) are also shown. Both models in most cases reproduce well the observed trend of annual runoff changes. Calculated increase of annual discharge at the outlet of the Lena River is around 748 $m^3$/s and 581 $m^3$/s per decade for ECOMAG and SWAP models, respectively (in other words, 235.9 $km^3$/decade and 183.2 $km^3$/decade, respectively). The observational data for 1979-2009 result in the increase of approximately 1000 $m^3$/s per decade (315.4 $km^3$/decade). The simulated ensemble mean Lena River discharge is statistically different from zero indicating that a considerable portion of the observed trend can be externally driven. For the Northern Dvina River, the simulated trends are insignificant, as well as the observed trend.

## 6. Conclusions

We have presented an analysis of large-basin hydrological response uncertainty originating from internal atmospheric variability that was for the first time performed with such a large (45 members) ensemble of climate model simulations. Internal variability is considered as one of three main factors of uncertainty in hydrological response to climate change (together with so-called “forcing” and “climate model” uncertainties). Importantly, in the presented simulations, the role of internal atmospheric variability is most visible for the time scales from years to first decades and for the regional spatial scales (e.g. Hawkins and Sutton, 2009), i.e. over spatial-temporal scales of water management in large river basins.

Our study focused on transformation of internal atmospheric variability by physically based hydrological models ECOMAG and SWAP and on impact of the variability on simulated runoff for the large Lena and Northen Dvina River basins located within the Arctic basin. It is important to emphasize, that due to stochastic nature of atmospheric variability, hydrological models driven by the output of a climate model are confined, as well as a climate model, to making projections rather than predictions (even in the past, not to mention the future), i.e. hydrological models are only able to provide information on statistical characteristics of runoff time series.

Internal atmospheric variability was simulated using ensemble simulations with the ECHAM5 atmospheric general circulation model. The ensemble consists of 45 simulations performed under identical prescribed lower boundary conditions (observed the sea surface temperature and the sea ice concentration for 1979-2012) and constant external forcing parameters corresponded to modern climate conditions. The only differences between the simulations were initial conditions of the atmosphere prescribed as instant atmospheric states changed by small perturbations.

The ensemble of the bias-corrected ECHAM5-outputs was assigned as distributed input for ECOMAG and SWAP hydrological models, and corresponding ensembles of runoff hydrographs were calculated for the Lena River and the Northern Dvina River. From these hydrographs, hydrological indicators, namely, the mean and the standard deviations of the annual, monthly and daily runoff, annual runoff trend, were estimated. Uncertainties of the hydrological indicators caused by the internal variability of the atmosphere were determined as normalized confidence intervals of the corresponding estimates.

The main findings of our research are the following:

1. Uncertainty in estimates of both the mean and the standard runoff deviation values increases with decreasing time interval of runoff averaging: from minimal uncertainty for annual runoff to maximal one for daily runoff. The mean annual runoff uncertainty originated from the internal variability of the atmosphere was found to be 6-10% depending on the model used and the study basin.

2. Atmospheric variability impact on uncertainties of the mean and the standard runoff deviation has a significant seasonal dependence. Uncertainties of monthly and daily runoff reach their maximum values during the periods of spring-summer snowmelt and summer-autumn rainfall floods for both rivers. Possible explanation of this finding is that physical mechanisms of flood events are more sensitive to intra-ensemble changes of the climate model outputs than more inertial mechanisms of low flow generation.

3. Simulated hydrographs for the Northern Dvina runoff are found to be more sensitive to internal atmospheric variability than those for the Lena River runoff. This is also manifested by the findings that runoff estimate uncertainties and their intra-annual irregularity are much higher for the Northern Dvina River simulations, when using both hydrological models. It is shown that increased effect of the internal atmospheric

variability in uncertainty of the Northern Dvina River runoff estimates can be explained by stronger non-linearity of runoff generation mechanisms compared to those of the Lena River basin.

4. Individual realizations of the simulated annual discharge series differ and are, in general, insignificantly correlated with the corresponding observed time series for both the Lena and the Northern Dvina River. However, for some individual realizations the linear link to observations is found to be quite strong: maximum correlation coefficients are 0.56 and 0.30 for the Lena and the Northern Dvina River simulations respectively.
5. It is shown that the averaging over large ensemble members effectively filters stochastic term related to internal atmospheric variability and results in an estimate of an externally forced signal related, in our experiments, to global sea surface temperature and sea ice concentration changes. We found that both models for the ensemble mean results reproduce the observed trend of the annual Lena River discharge. The simulated trends are (close to) normally distributed around the ensemble mean value that indicates, for the Lena River discharge, that a considerable portion of the observed trend can be externally driven. The trend for the Northern Dvina River changes appeared to be insignificant both for the simulation results and the observational data. This assumes a dominant role of internal variability in generating the Northern Dvina runoff changes during the simulation period.

Our results, in line with the conclusions of Deser et al. (2012; 2014) who analyzed temperature and precipitation changes, assume the importance of performing large ensembles of climate change projections with climate models also for making robust estimates of uncertainty and externally forced signal in hydrological response on decadal to multi-decadal time scale.

## Acknowledgements

We express our sincere gratitude to Dr. Stefan Hagemann and an anonymous reviewer for constructive comments that helped to improve the manuscript. The presented researches related to the Northern Dvina River hydrological modeling and simulation results analysis were financially supported by the Russian Science Foundation (grant No. 14-17-00700). The part of the researches related to analyzing climate model experiments and the Lena River hydrological simulations was financially supported by the Russian Ministry of Education and Science (grant No. 14.B25.31.0026). ECHAM5 climate model simulations were performed at the Northern Germany High Performance Computing Center (HLRN) supported by EU NACLIM project.

The present work has been carried out within the framework of the Panta Rhei Research Initiative of the International Association of Hydrological Sciences (IAHS).

## References

Anagnostopoulos, G. G., Koutsoyiannis, D., Christofides, A., Efstratiadis, A. and Mamassis, N.: A comparison of local and aggregated climate model outputs with observed data, Hydrol. Sci. J., 55(7), 1094–1110, 2010.

Bates, B. C., Kundzewicz, Z. W., Wu, S. and Palutikof, J. P. (eds): Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, 2008.

Braun, M., Caya, D., Frigon, A. and Slivitzky, M.: Internal variability of Canadian RCM's hydrological variables at the basin scale in Quebec and Labrador, J. Hydrometeorol., 13, 443–462, 2012.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: a review, Hydrol. Processes, 9, 251–290, 1995.

Chiew, F.H.S., Teng, J., Vaze, J., Post, D.A., Perraud, J.M. and Kirono, D.G.C. and Viney, N.R..: Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modelling method, Water Resources Research, 45, W10414, doi:10.1029/2008WR007338, 2009.

Deque, M., D.P. Rowell, D. Luethi, F. Giorgi, J.H. Christensen, B. Rockel, D. Jacob, E. Kjellstroem, M. de Castro and van den Hurk, B.J.J.M.: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections, Climatic Change, 81, 53-70, 2007.

Deser, C., Phillips, A., Bourdette, V. and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Clim. Dynam., 38, 527–546, 2012.

Deser, C., Phillips, A.S., Alexander, M.A., and Smoliak, B.V.: Projecting North American Climate over the Next 50 Years: Uncertainty due to Internal Variability, J. Climate, 27, 2271–2296, 2014.

Driessen, T.L.A., Hurkmans, R.T.W.L., Terink, W., Hazenberg, P., Torfs, P.J.J.F., and Uijlenhoet, R.: The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model, HESS, 14, 651–665, 2010.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, USA, 2013.

Gelfan, A.N. and Motovilov, Yu.G.: Long–term hydrological forecasting in cold regions: retrospect, current status and prospect, Georgaphy Compass, 3(5), 1841–1864, 2009.

Gusev, Ye.M. and Nasonova, O.N.: The Land Surface Parameterization scheme SWAP: description and partial validation, Global and Planetary Change, 19(1-4), 63-86, 1998.

Gusev, Ye.M., Nasonova, O.N.: Modelling heat and water exchange in the boreal spruce forest by the land-surface model SWAP, J. Hydrology, 280(1-4), 162-191, 2003.

Gusev, E.M., Nasonova, O.N., Dzhogan, L.Ya. and Kovalev, E.E.: Northern Dvina runoff simulation using land-surface model SWAP and global databases, Water Res, 38(4), 470-483, 2011.

Gusev, E.M., Nasonova, O.N. and Dzhogan, L.Ya.: Physically-based modelling long-term dynamics of daily river streamflow and snow cover in the Lena River Basin, Water Res., 2015 (in press).

Hagemann. S., Arpe, K., and Roeckner, E.: Evaluation of the hydrological cycle in the ECHAM5 model, J. Climate, 19, 3810-3827, 2006.

Hagemann, S., H. Göttel, D. Jacob, P. Lorenz and Roeckner, E.: Improved regional scale processes reflected in projected hydrological changes over large European catchments. Climate Dynamics 32 (6), 767-781, 2009.

Hagemann, S., Loew, A., Andersson, A.: Combined evaluation of MPI-ESM land surface water and energy fluxes, J. Adv. Model. Earth Syst., 5, 259-286, 2013.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, B. Am. Meteorol. Soc., 90:1095, doi:10.1175/2009BAMS2607.1, 2009.

Klemes, V.: Operational testing of hydrological simulation models. Hydrological Sciences Journal, 31, 13–24, 1986.

Koutsoyiannis, D., Montanari, A., Lins, H.F. and Cohn, T.A.: Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research, Hydrological Sciences Journal, 54(2), 394-405, 2009.

Krylenko, I., Motovilov, Yu., Antokhina, E., Ghuk, V. and Surkova, G.: Physically based distributed modelling of river runoff under changing climate conditions, in: Remote Sensing and GIS for Hydrology and Water Resources (Proceedings RSHS14 and ICGRHWE14), Guangzhou, China, August 2014.

Kundzewicz, Z.W., Mata, L.J., Arnell, N.W., Döll, P., Jimenez, B., Miller, K., Oki, T., Şen, Z. and Shiklomanov, I.: The implications of projected climate change for freshwater resources and their management, Hydrological Sciences Journal, 53 (1), 3–10, 2008.

Kundzewicz, Z.W.and Stakhiv, E.Z.: Are climate models “ready for prime time” in water resources management applications, or is more research needed?, Hydrological Sciences Journal, 55(7), 1085-1089, 2010.

Lafaysse, M., B. Hingray, A. Mezghani, J. Gailhard and L. Terray: Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin, Water Resour. Res., 50, 3317–3341, doi:10.1002/2013WR014897, 2014.

Lins, H.F.: Observed Trends in Hydrologic Cycle Components. In: Encyclopedia of Hydrological Sciences, Edited by MG Anderson, 2005.

Latif, M. and Keenlyside, N.S.: A perspective on decadal climate variability and predictability, Deep-Sea Research II, 58, 1880–1894, 2011.

Latif, M., Martin, T. and Park, W.: Southern Ocean Sector Centennial Climate Variability and Recent Decadal Trends, J. Climate, 26, 7767-7782, 2013.

Lorenz, Edward N.: Deterministic Nonperiodic Flow, Journal of the Atmospheric Sciences, 20(2), 130–141, 1963.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E.: The WCRP CMIP3 multimodel dataset – A new era in climate change research, B. Am. Meteorol. Soc., 88, 1383–1394, doi:10.1175/bams-88-9-1383, 2007.

Motovilov, Yu.G. and Gelfan, A.N.: Assessing runoff sensitivity to climate change in the Arctic basin: empirical and modelling approaches. In: Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections A. Gelfan, D. Yang, E. Gusev and H. Kunstmann (eds), IAHS Publications 360, 105-112, 2013.

Motovilov, Yu., Gottschalk, L., Engeland, K., and Belokurov, A.: ECOMAG – regional model of hydrological cycle. Application to the NOPEX region. Department of Geophysics, University of Oslo, Institute Report Series no. 105. 1999a.

Motovilov Yu., Gottschalk L., Engeland K. and Rodhe A.: Validation of a distributed hydrological model against spatial observation, Agricultural and Forest Meteorology, 98-99, 257-277, 1999b.

Nash, I. E. and Sutcliffe, I.V.: River flow forecasting through conceptual models. Journal of Hydrology, 10, 282–290, 1970.

Oki, T., Nishimura, T. and Dirmeyer P.: Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP), J. Meteorol. Soc. of Japan., 77, 1B, 235 – 255, 1999.

Praskievicz, S. and Chang, H.: A review of hydrological modelling of basin-scale climate change and urban development impacts, Progress in Physical Geography, 33(5), 650–671, 2009.

Peel, M.C. and Blöschl, G.: Hydrological modelling in a changing world, Progress in Physical Geography, 35(249), doi: 10.1177/0309133311402550, 2011.

Peel, M. C., Srikanthan, R., McMahon, T. A. and Karoly, D. J.: Uncertainty in runoff based on Global Climate Model precipitation and temperature data – Part 2: Estimation and uncertainty in annual runoff and reservoir yield, Hydrol. Earth Syst. Sci. Discuss., 11, 4579–4638, doi:10.5194/hessd-11-4579-2014, 2014.

Räisänen, J.: CO2-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability, J. Climate, 14, 2088–2104, 2001.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res., 108, 4407, doi:10.1029/2002JD002670, D14, 2003.

Refsgaard, J.C, Madsen, H, Andréassian, V, Arnbjerg-Nielsen, K, Davidson, T.A, Drews, M, Hamilton, D.P, Jeppesen, E, Kjellström, E, Olesen, J.E, Sonnenborg, T.O, Trolle, D, Willems, P, and Christensen, J.H.: A framework for testing the ability of models to project climate change and its impacts. Climatic Change, 122, 271-282., 10.1007/s10584-013-0990-2, 2014

Roeckner, E., G. Bäuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, A. Rhodin, U. Schlese, U. Schulzweida, and Tompkins, A., The atmospheric general circulation model ECHAM5. Part I: Model description. Max Planck Institute for Meteorology Rep. 349, 127 pp., 2003.

Sankarasubramanian, A., Vogel, R.M. and Limbrunner, J.F.: Climate elasticity of streamflow in the United States, Water Resources Research, 37(6), 1771–1781, 2001.

Sansom, P. G., Stephenson, D. B., Ferro, C. A. T., Zappa, G. and Shaffrey, L.: Simple Uncertainty Frameworks for Selecting Weighting Schemes and Interpreting Multimodel Ensemble Climate Change Experiments, J. Climate, 26 (12), 4017–4037, doi:10.1175/JCLI-D-12-00462.1, 2013.

Seiller, G. and Anctil, F.: Climate change impacts on the hydrologic regime of a Canadian river: comparing uncertainties arising from climate natural variability and lumped hydrological model structures, Hydrol. Earth Syst. Sci., 18, 2033–2047, doi:10.5194/hess-18-2033-2014, 2014.

Semenov, V.A.: Role of sea ice in formation of wintertime Arctic temperature anomalies, Izvestiya, Atmospheric and Oceanic Physics, 50(4), 343–349, 2014.

Semenov, V.A., Latif, M., Dommenget, D., Keenlyside, N.S., Strehz, A., Martin, T. and Park, W.: The Impact of North Atlantic-Arctic Multidecadal Variability on Northern Hemisphere Surface Air Temperature, J. Climate, 23, 5668-5677, 2010.

Semenov, V.A. and Latif, M.: The early twentieth century warming and winter Arctic sea ice, The Cryosphere, 6, 1231–1237, 2012.

Singh, R., Wagener, T., van Werkhoven, K., Mann, M. E. and Crane, R.: A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate – accounting for changing watershed behaviour, Hydrol. Earth Syst. Sci., 15, 3591-3603, doi:10.5194/hess-15-3591-2011, 2011.

Taylor, K.E., Stouffer, R. J. and Meehl, G. A.: An Overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485-498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Teutschbein, C. and Seibert, J.: Regional Climate Models for Hydrological Impact Studies at the Catchment Scale: A Review of Recent Modeling Strategies, Geography Compass, 4, 834–860, doi:10.1111/j.1749-8198.2010.00357.x, 2010.

Vano, J.A. and Lettenmaier, D. P.: A sensitivity-based approach to evaluating future changes in Colorado River discharge. Climatic Change 122, 621-634, 2014.

Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R. and Turcotte, R.: An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources, Hydrol. Earth Syst. Sci., 17, 565-578, 2013.

Wilby, R.L.: Evaluating climate model outputs for hydrological applications – Opinion, Hydrol. Sci. J., 55(7), 1090–1093, 2010.

Yip, S., Ferro, C. A. T., Stephenson, D. B. and Hawkins, E.: A simple, coherent framework for partitioning uncertainty in climate predictions, J. Climate, 24(17), 4634–4643, doi:10.1175/2011JCLI4085.1, 2011.

Table 1 The Nash and Sutcliffe efficiency, NSE, and bias evaluation criteria calculated from simulated and measured daily discharge at the outlets of Lena and Northern Dvina River basins

| River (Gauge) | Period | NSE | Bias, % |
|---|---|---|---|
| ECOMAG (calibration period) | | | |
| Lena (Stolb) | 2000-2009 | 0.90 | -2.9 |
| N. Dvina(Ust'-Penega) | 2000-2009 | 0.88 | 1.4 |
| ECOMAG (validation period) | | | |
| Lena (Stolb) | 1987-1999 | 0.86 | 1.4 |
| N. Dvina(Ust'-Penega) | 1970-1999 | 0.81 | 2.0 |
| SWAP (calibration period) | | | |
| Lena (Stolb) | 1971-1977 | 0.82 | -4.9 |
| N. Dvina(Ust'-Penega) | 1986-1990 | 0.86 | -1.1 |
| SWAP (validation period) | | | |
| Lena (Stolb) | 1978-1999 | 0.80 | -3.7 |
| N. Dvina(Ust'-Penega) | 1967-1985; 1991-1998 | 0.85 | -0.6 |

Table 2. Uncertainty indices $UN(M)$ (in %) for $M$-estimates of annual, monthly and daily runoff

| Runoff characteristic | Lena River | | Northern Dvina River | |
|---|---|---|---|---|
| | ECOMAG | SWAP | ECOMAG | SWAP |
| Annual runoff | 6 | 7 | 10 | 7 |
| Monthly runoff | 7 | 11 | 19 | 19 |
| January | 3 | 9 | 5 | 9 |
| February | 2 | 8 | 2 | 9 |
| March | 1 | 8 | 5 | 23 |
| April | 1 | 24 | 33 | 41 |
| May | 21 | 9 | 10 | 23 |
| June | 6 | 9 | 14 | 18 |
| July | 8 | 9 | 22 | 9 |
| August | 10 | 9 | 32 | 14 |
| September | 13 | 10 | 35 | 17 |
| October | 10 | 11 | 29 | 21 |
| November | 8 | 12 | 22 | 24 |
| December | 5 | 13 | 17 | 19 |
| Daily runoff | 8 | 12 | 24 | 21 |

Table 3. Uncertainty Indices $UN(SD)$ (in %) for $SD$-estimates of the Annual and Monthly runoff

| Runoff characteristic | Lena River | | Northern Dvina River | |
|---|---|---|---|---|
| | ECOMAG | SWAP | ECOMAG | SWAP |
| Annual runoff | 24 | 26 | 30 | 31 |
| Monthly runoff | 32 | 30 | 52 | 33 |
| January | 29 | 35 | 85 | 29 |
| February | 30 | 33 | 95 | 29 |
| March | 30 | 25 | 104 | 36 |
| April | 31 | 23 | 36 | 42 |
| May | 84 | 55 | 24 | 45 |
| June | 25 | 21 | 27 | 39 |
| July | 29 | 17 | 39 | 23 |
| August | 25 | 26 | 46 | 26 |
| September | 26 | 28 | 47 | 27 |
| October | 22 | 34 | 37 | 30 |
| November | 23 | 32 | 33 | 29 |
| December | 28 | 33 | 51 | 35 |
| Daily runoff | 45 | 36 | 98 | 45 |

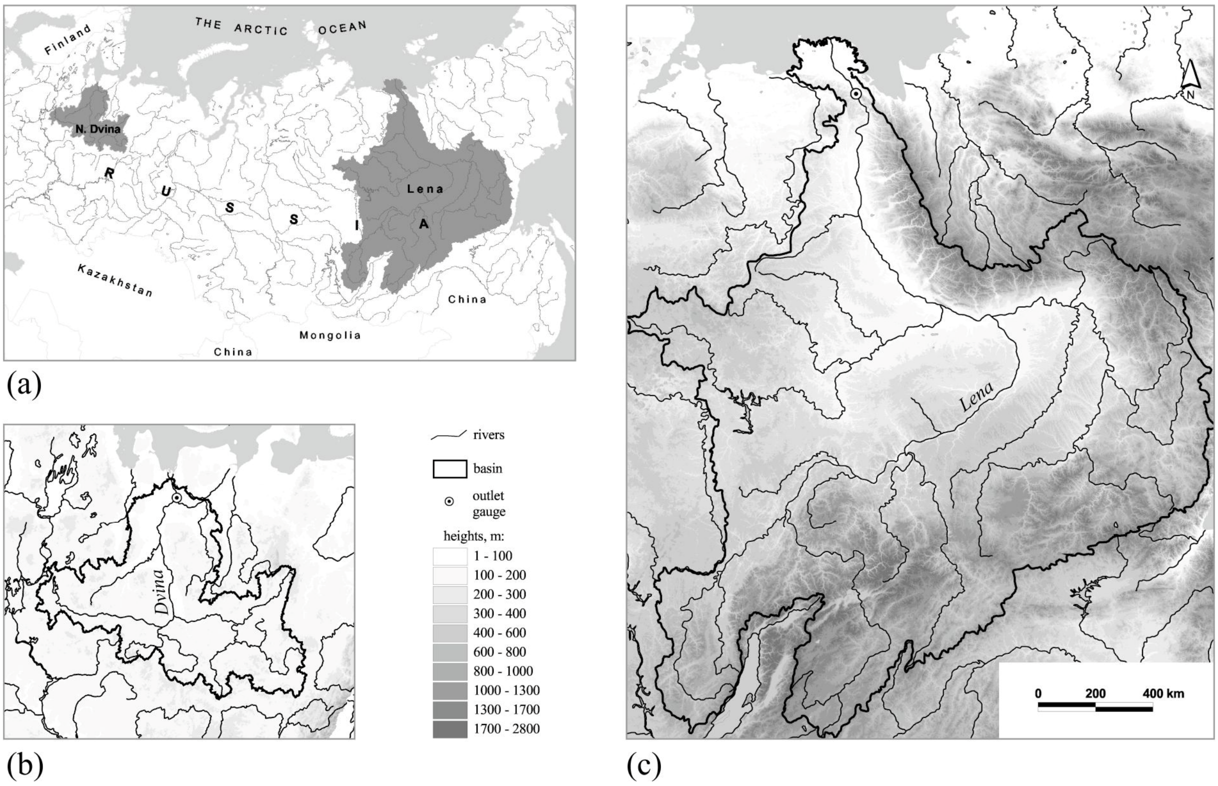


Figure 1. Case study basins: location (a), Northern Dvina River basin (b), Lena River basin (c)

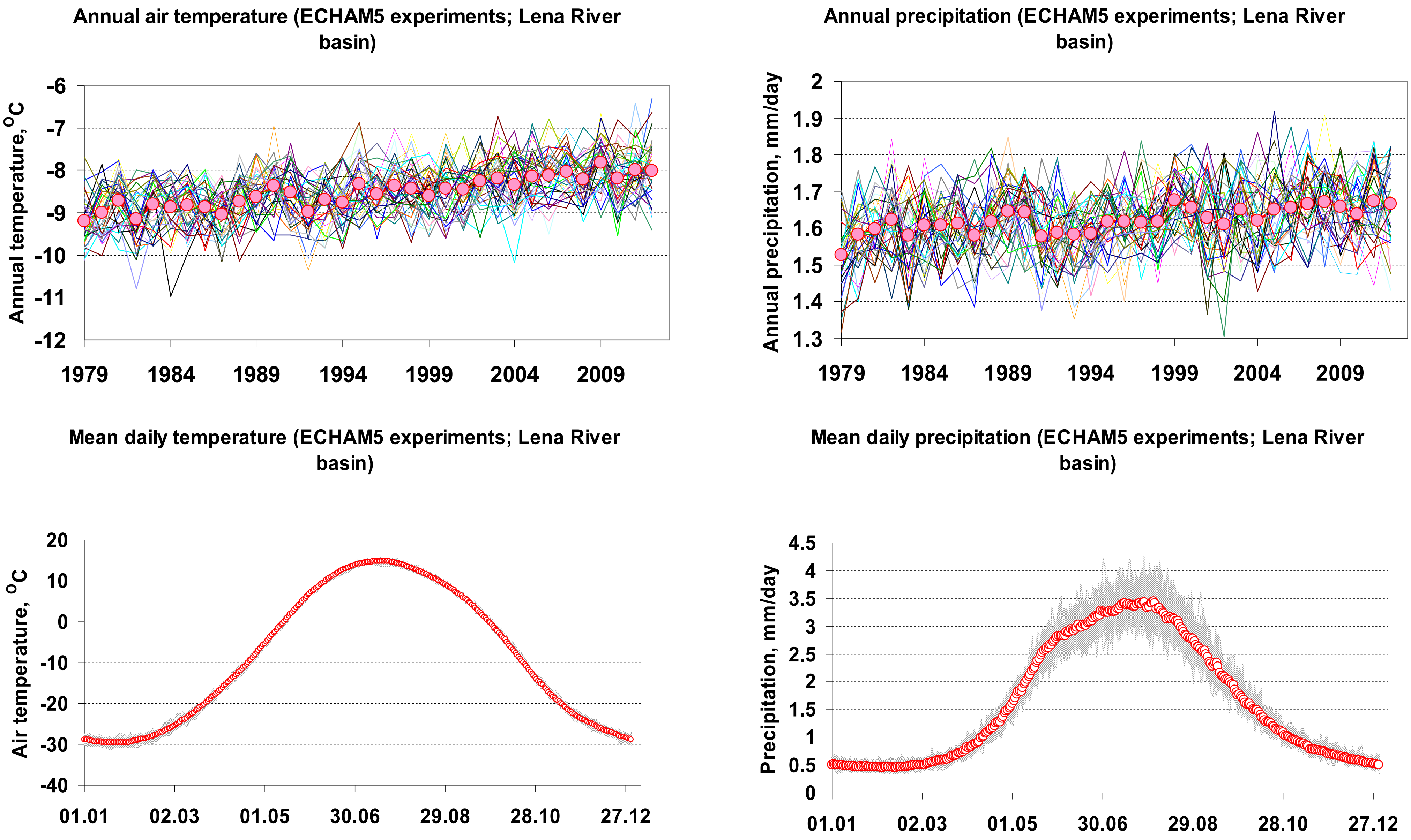


Figure 2. ECHAM5-simulated ensembles of mean annual surface air temperature (SAT) (top; left) and precipitation (top; right), as well as mean daily SAT (bottom; left) and precipitation (bottom; right) averaged over the Lena River basin. Dots in top figures and bold line in bottom figures denote corresponding ensemble mean values

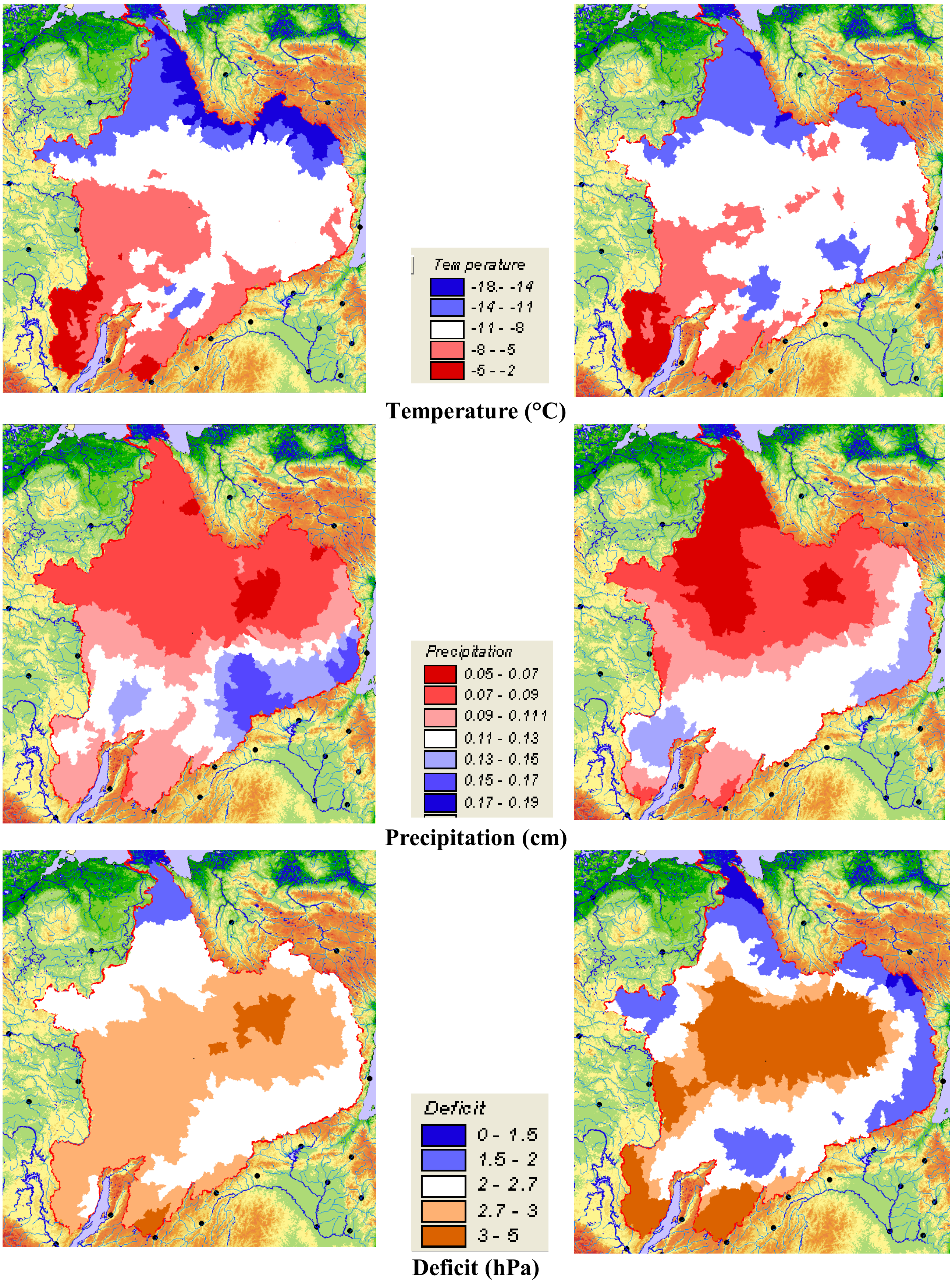


Figure 3. Observed (left) and the bias-corrected ECHAM5-simulated (right) patterns of mean annual values of air temperature (°C), precipitation (cm) and air humidity deficit (hPa) within the Lena River basin.

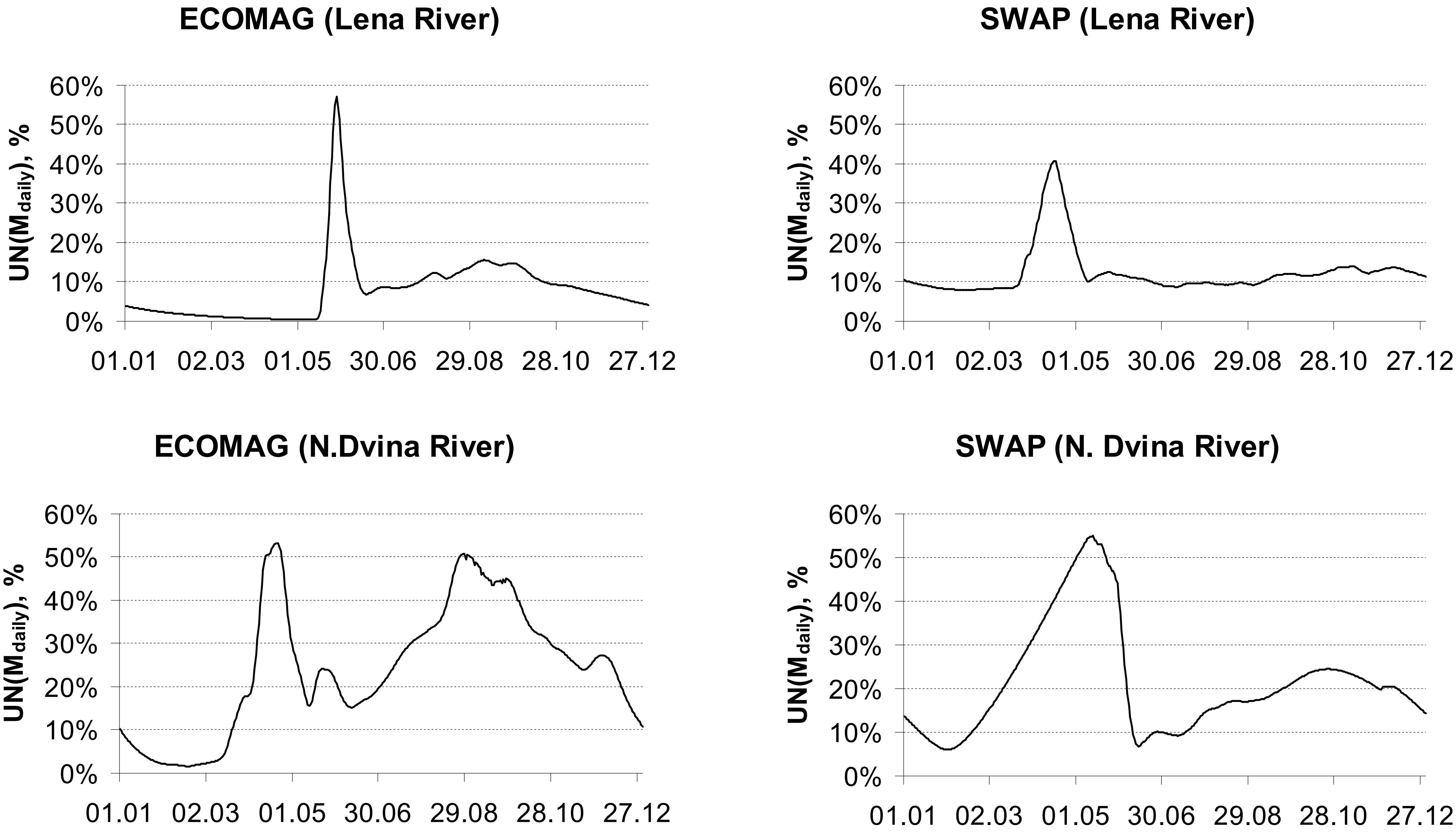


Figure 4. Intra-annual variation of uncertainty indices $UN(M)$ (in %) for the $M$ -estimates of daily runoff

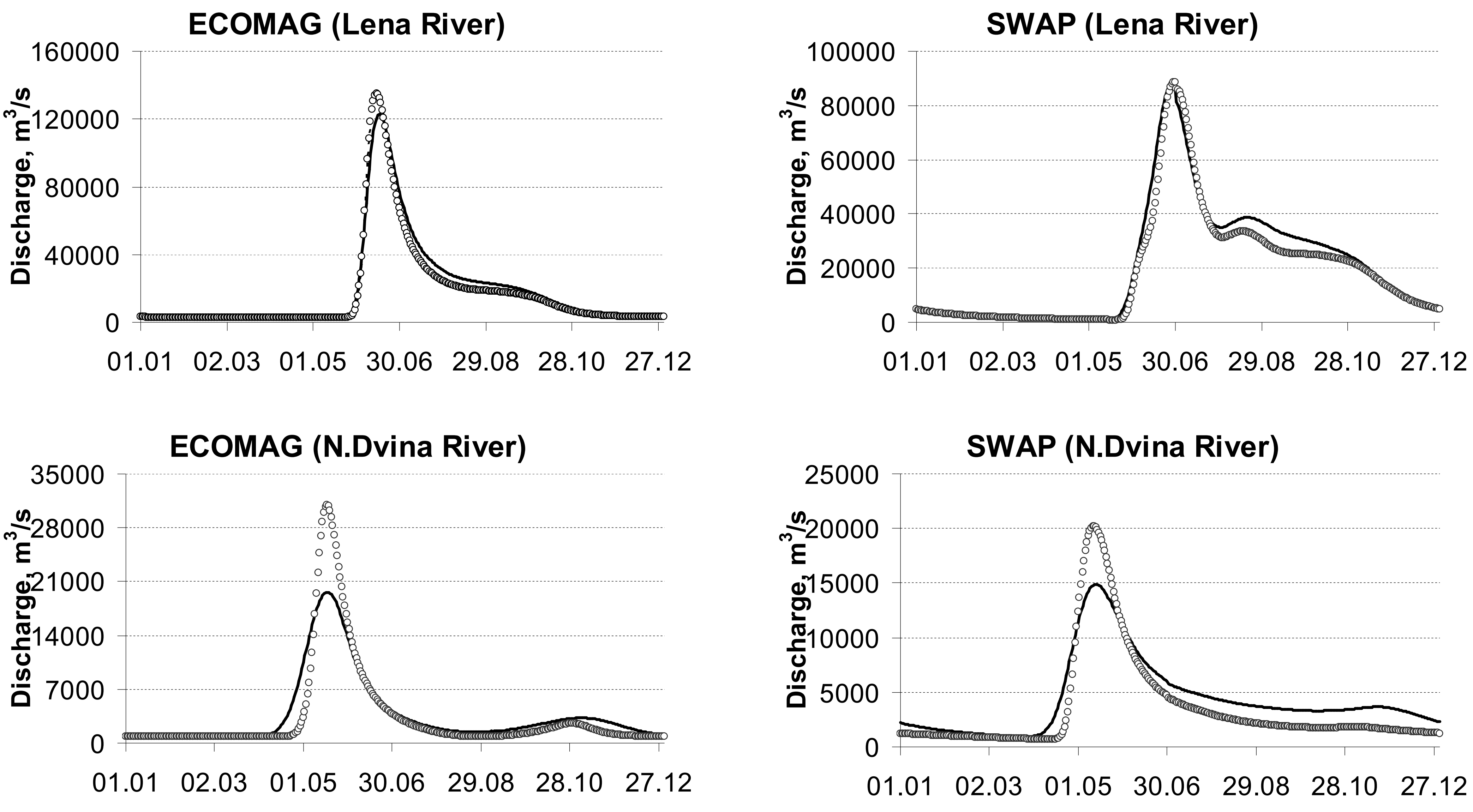


Figure 5. Mean hydrographs calculated as an averaged response to ensemble input (solid line) and as a response to ensemble averaged input (dotted line)

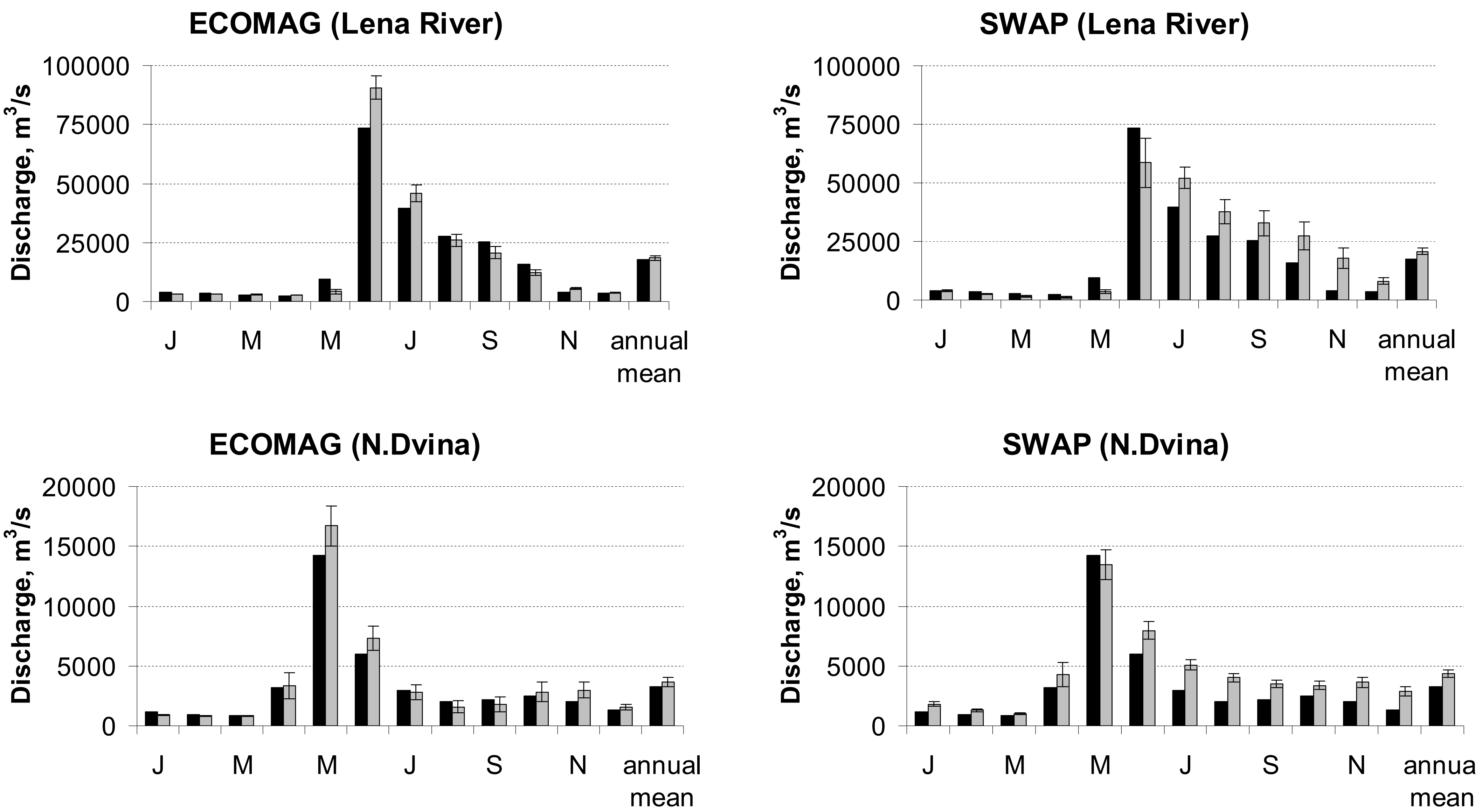


Figure 6. $M$ -estimates of annual and monthly discharges at the outlets of the Lena River (top) and the Northern Dvina River (bottom).

- black columns show estimates obtained from the observation data for 1979-2009.

- gray columns show estimates obtained from the ensemble simulation (with indicated 95% confidence intervals $\gamma_M$ for these estimates)

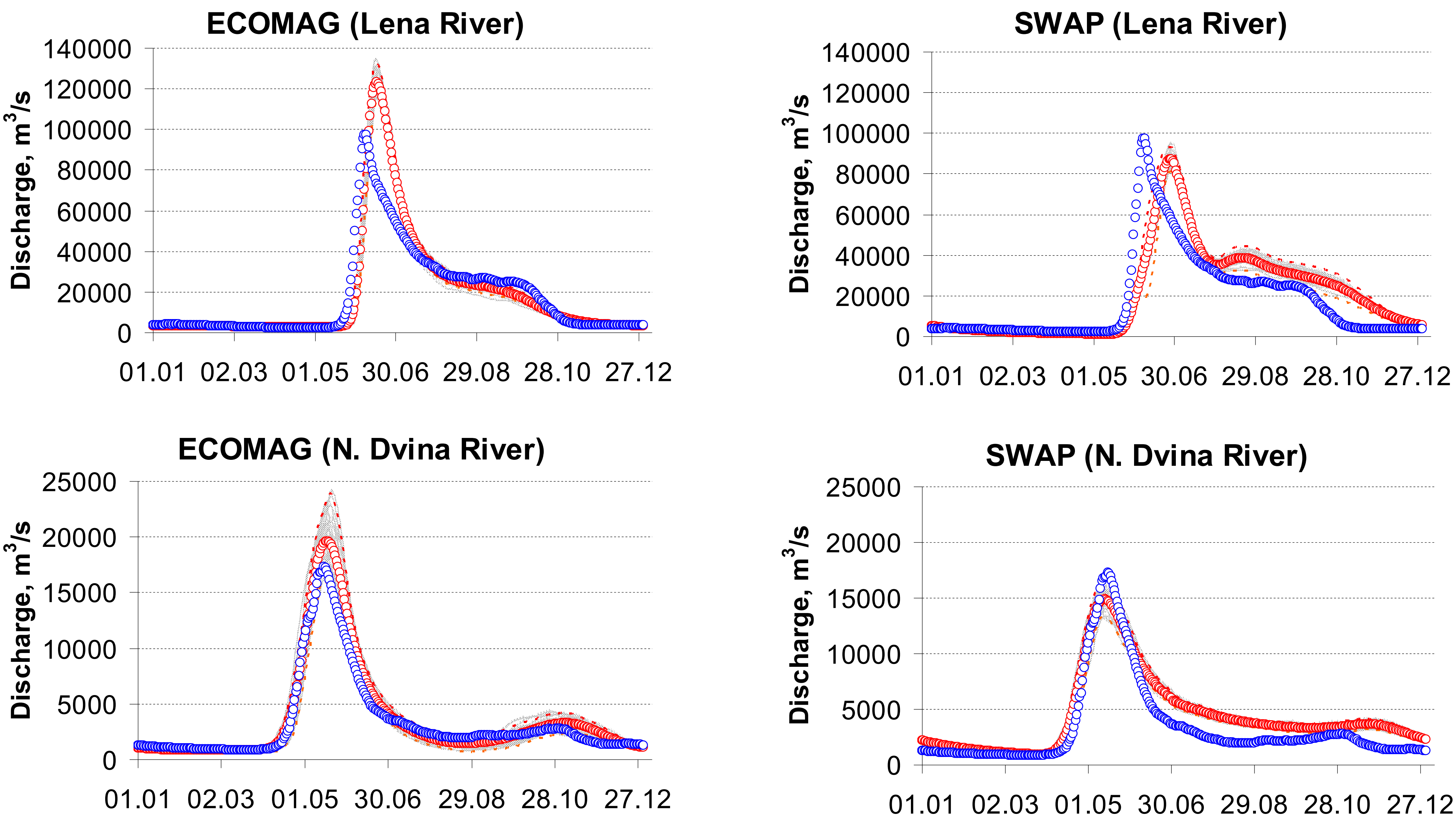


Figure 7. $M$ -estimates of the daily discharges at the outlets of the Lena River (top) and the Northern Dvina River (bottom)

- blue points show estimates based on observational data for the period of 1979-2012.

- red points show estimates based on ensemble simulations (gray thin lines).

- red dotted line shows the boundaries of 95% confidence interval of mean daily discharges.

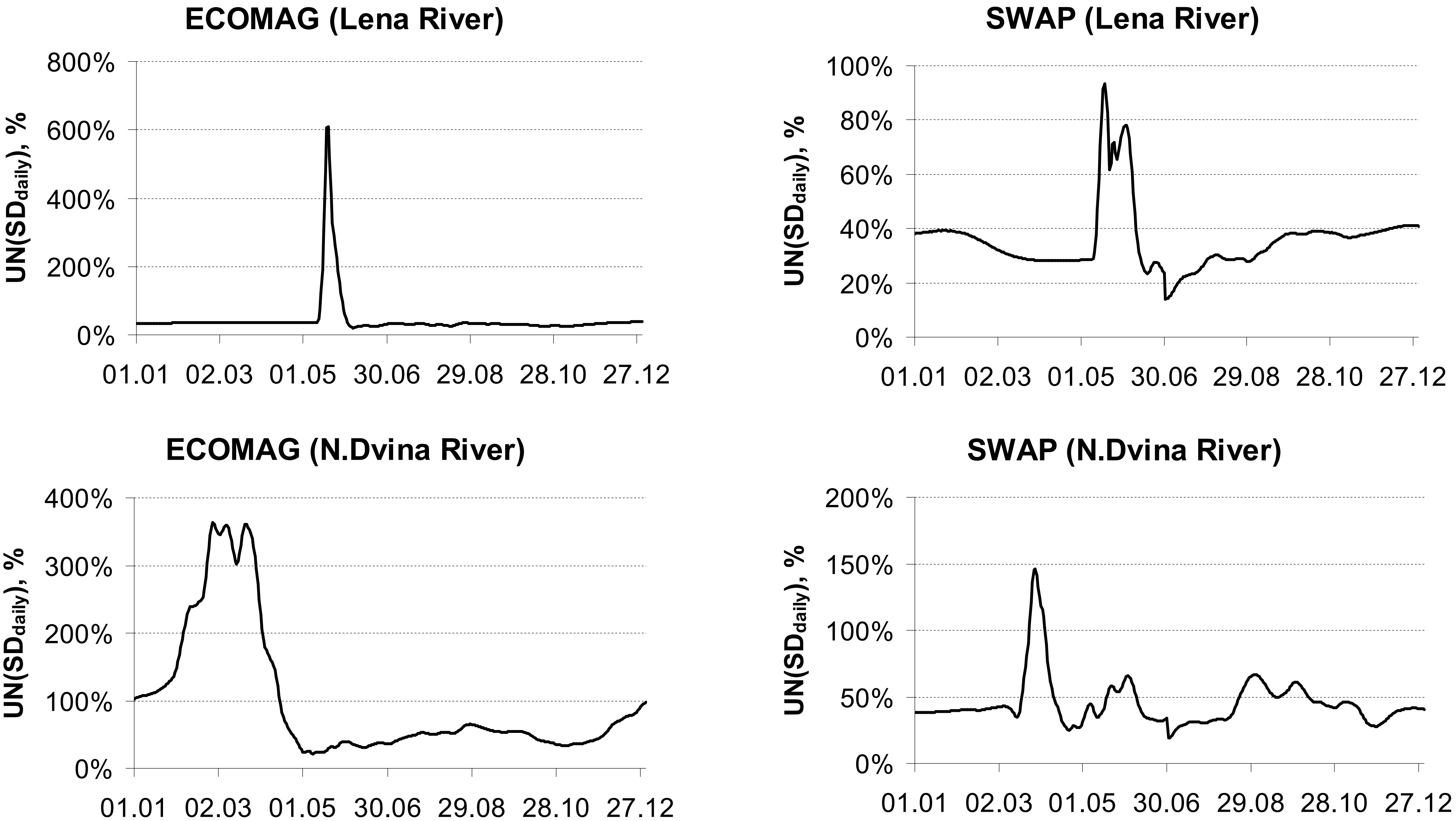


Figure 8. Uncertainty indices $UN(SD)$ (in %) for the $SD$-estimates of the daily runoff

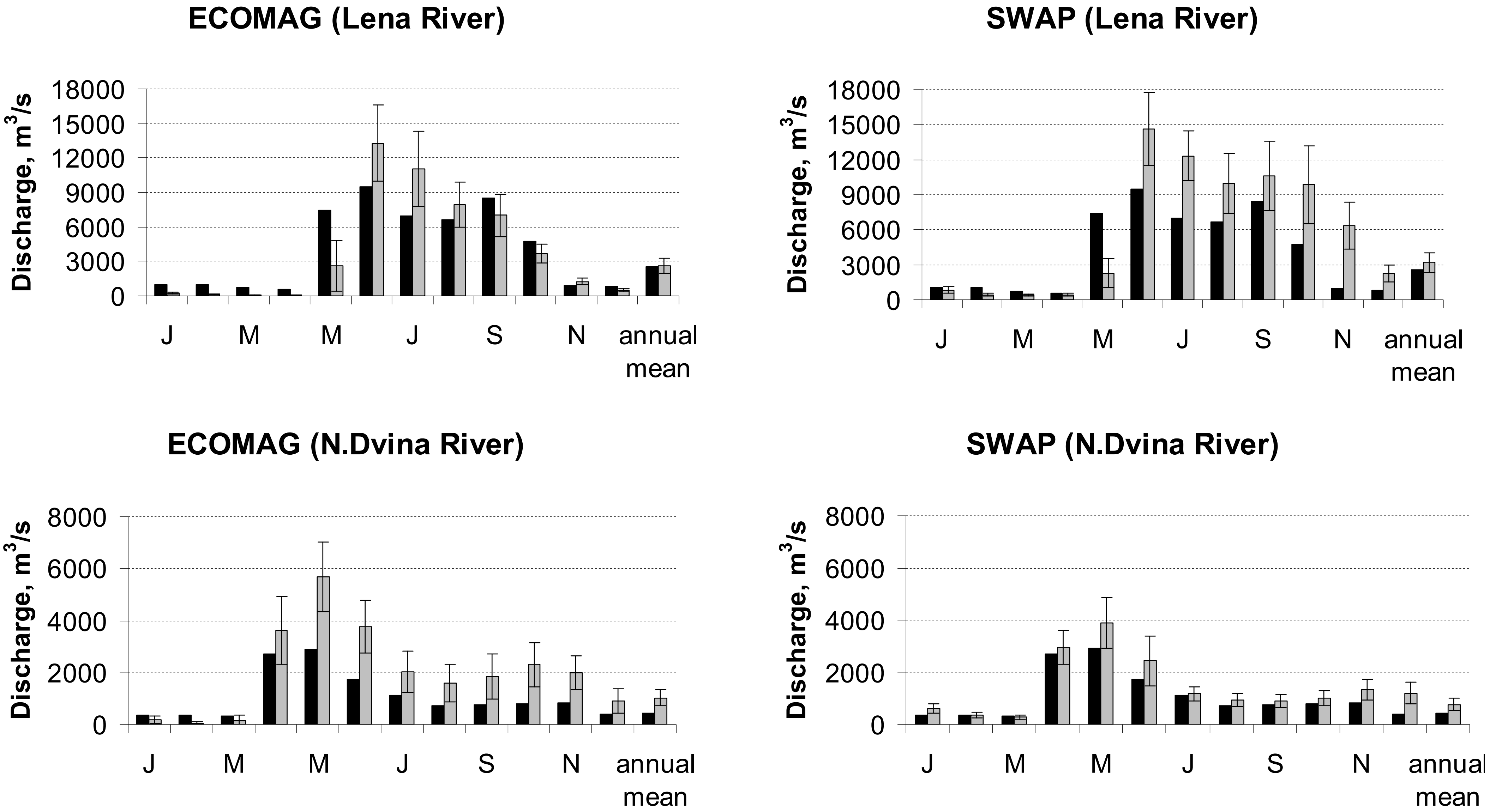


Figure 9. *SD* -estimates of the annual and monthly discharges at the outlets of the Lena River (top) and the Northern Dvina River (bottom).

- black columns show estimates obtained from the observational data for 1979-2009.

- gray columns show estimates obtained from the ensemble simulation (with indicated 95% confidence intervals $\gamma_{SD}$ for these estimates)

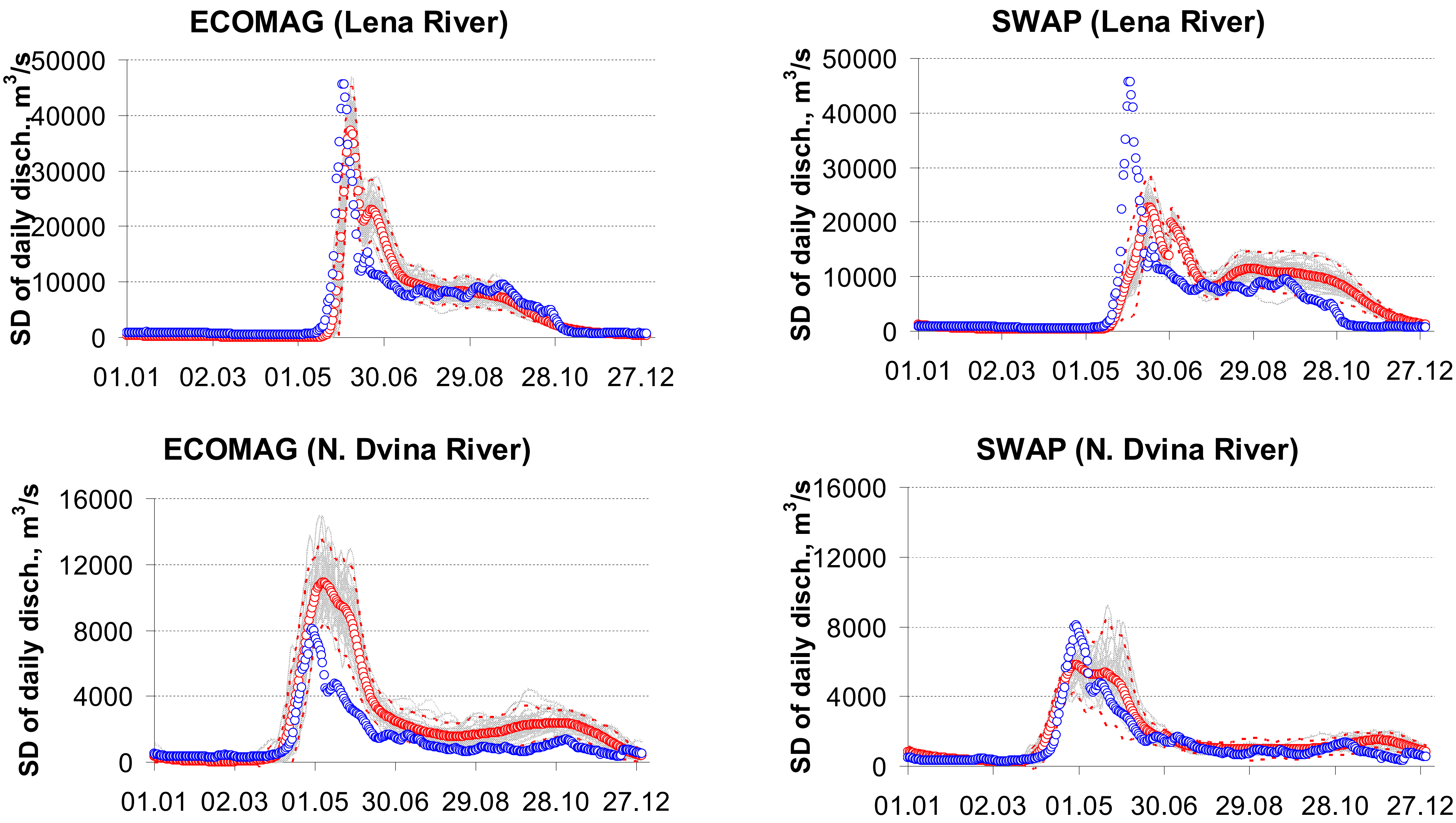


Figure 10. *SD* -estimates of the daily discharges at the outlets of the Lena River (top) and the Northern Dvina River (bottom)

- blue points show estimates based on observational data for the period of 1979-2012.

- red points show estimates based on ensemble simulations (gray thin lines).

- red dotted line shows the boundaries of 95% confidence interval of mean daily discharges

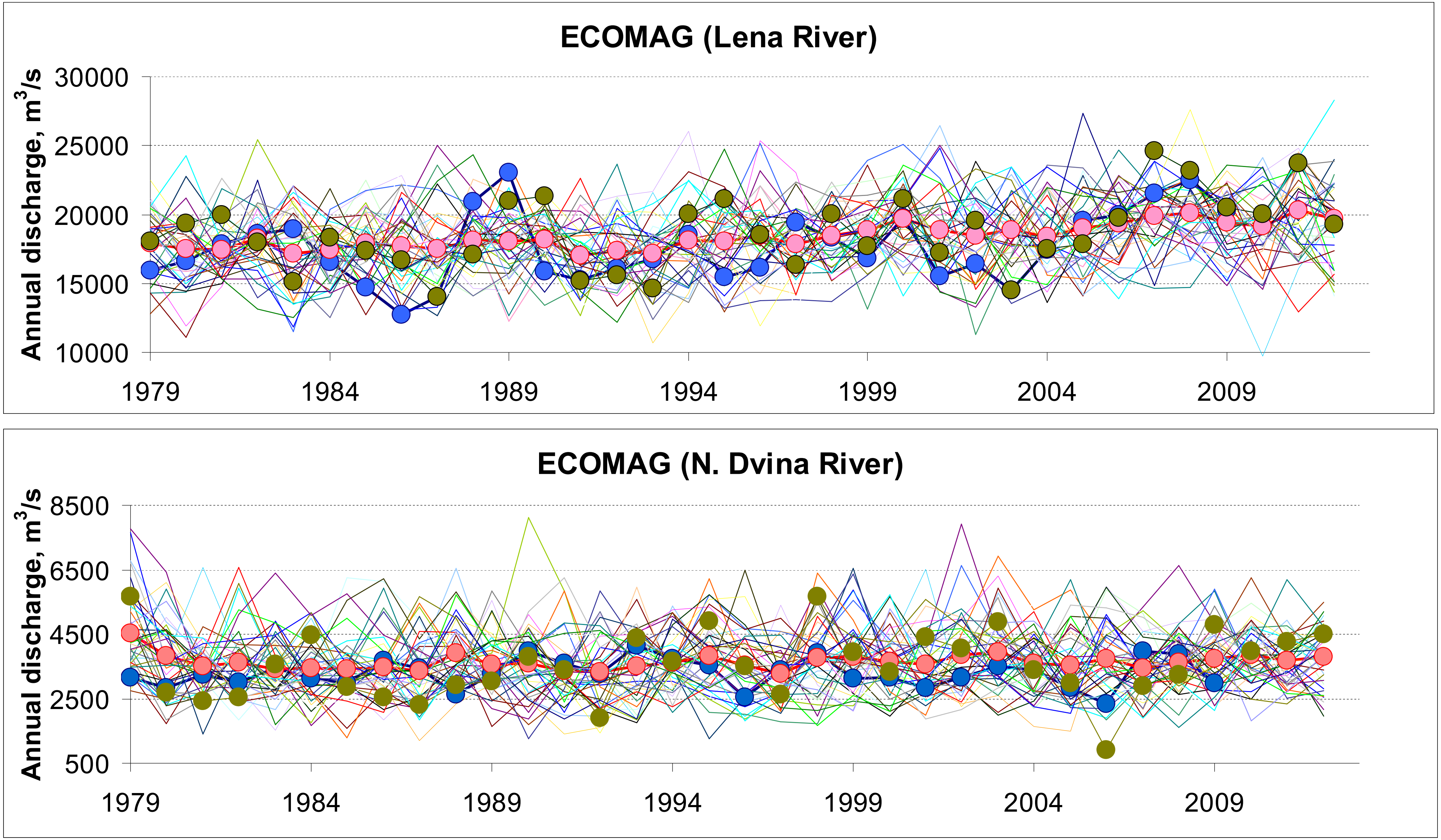


Figure 11. Observed (line with blue markers) and simulated series of annual discharges

- thin lines show ensemble (45 realizations) of the calculated annual discharges

- the line with red markers shows the ensemble mean

- the line with green markers shows the realization most strongly correlated with the observed time series

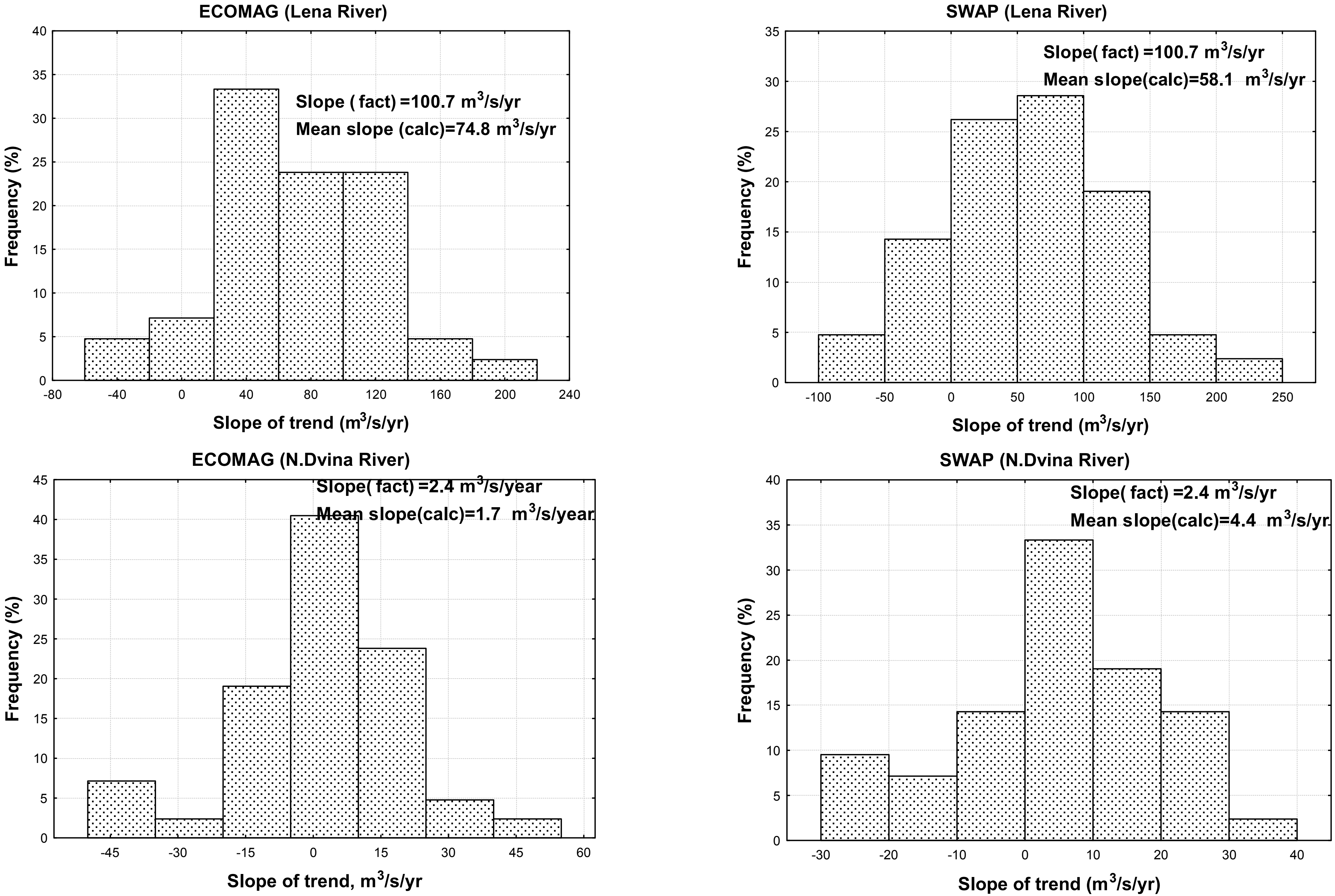


Fig. 12. Histograms of the linear trend slope derived from the ensembles of simulated annual discharge time series