ELSEVIER

Regression-based downscaling of spatial variability for hydrologic applications

G. Bürger[a,]*, Y. Chen[b]

[a]*Fu Berlin, Institüt für Meteorologie, Carl-Heinrich-Becker-Weg 6–10, 12165 Berlin, Germany*
[b]*Max-Planck-Institute für Biogeochemie, Hans Knöll Str. 10,07745 Jena, Germany*

Abstract

There is an obvious imbalance between, on the one hand, the importance of spatio-temporal variability of precipitation for river flows and, on the other, their representation in current empirical downscaling models that are applied for climate scenarios. The imperfect variability results from incomplete forcing of the large scales. The last IPCC report mentioned three regression-based methods that try to overcome the imperfection of point-wise variability: randomization, inflation, and expanded downscaling, Here, we analyze and compare these methods with respect to their spatial variability and how that relates to river runoff. Using the downscaled temperature and precipitation for observed and simulated large-scale forcings (climate scenarios), we applied the hydrologic model HBV for two river basins in Germany. We discuss the obvious and hidden model imperfections regarding present and future precipitation climate, along with their relevance for runoff. The overall picture is quite diverse, and it appears that temporal characteristics, i.e. time-lagged effects, are at least as important as spatial characteristics. We conclude that, although the models agree in a number of essential projections for river flow, a more consistent picture requires the full spatio-temporal variability as it depends on the large scale atmosphere.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Randomization; Inflation; Expanded downscaling; Hydrological scenarios; Spatial correlations

1. Introduction

River runoff, in its extreme forms of floods and droughts, is one of the most vitally felt possible impacts of global warming. And while at first sight runoff might appear as a handy and simple 1-dimensional quantity being easy to calibrate and validate, it is known to be among the most complex and challenging phenomena in the field of climate impact research. This is owed to the circumstance that runoff acts as a spatial and temporal integrator of meteorological fields, and is thus sensitive to the entire spatio-temporal structure of those fields. Inadequate representation of either the spatial correlations and/or the temporal autocorrelation of these fields leads to errors in the simulated runoff, which are thus rather complex and difficult to isolate.

Supposedly no other field therefore requires such a detailed knowledge about the spatio-temporal

* Corresponding author.
E-mail addresses: gerd.buerger@mail.met.fu-berlin.de
(G. Bürger), ychen@bgc-jena.mpg.de (Y. Chen).

signature of climate change than hydrologic systems. The simulation of climate change is done using general circulation models (GCMs) that are driven by estimates of historic and future greenhouse gas emissions, and which convey the global climate information on a fairly large-scale coordinate system. It is obvious that these models are incapable of giving any such detail. This information must be externally derived from procedures known as *downscaling*. However, downscaling often reveals a trade off between the spatial and temporal detail, in the sense that highly resolved temporal scales are often poor in the spatial domain, and vice versa. For example, daily precipitation scenarios are usually encountered as point simulations for a number of recording stations (Karl et al., 1990; Bardossy and Plate, 1991; Wilby and Wigley, 1997; Bronstert et al., 2000; Bürger, 2002), while spatially explicit simulations are preferably encountered in a context with a poorly resolved temporal scale (such as monthly or annually, cf. Grabs, 1997; Kilsby, 2000; Kleinn et al., 2002).

Through the event of regional circulation models (RCMs) a reconciliation between the two domains seems possible. But despite enormous progress with these models—the actual resolution is 1/6 degrees—RCM errors are still too large, especially over complex terrain, for a direct coupling to hydrologic models (Machenhauer et al., 1998; Giorgi and Mearns, 1999; Jacob et al., 2001).

If observations are not a limiting factor empirical downscaling techniques offer a very pragmatic alternative to the complex and expensive dynamical simulations, for hydrologic applications in particular, since the spatial and temporal detail of a dense observational network can be fully utilized. Many different schemes have been developed, all of them in some way or another formalizing the experience that local weather is governed, at least partly, by the prevailing large-scale circulation patterns. The main question is now: How does one deal with the unexplained part? Or how can one link the unexplained to the explained part? For hydrologic applications this is of major importance as, e.g. the daily variation of precipitation is largely unexplained but contributes strongly to the overall river flow variability. Since in this case usually no other information is available stochastic, or quasi-stochastic elements populate the downscaling schemes to restore the original local

variability. This can be done in various ways, each of which has its own advantage and disadvantage. Roughly, the available methods can be grouped into three different schemes, one using weather-types (cf. Bardossy and Plate, 1992; Charles et al., 1999; Wilks 1999), one using resampling techniques or analogues (cf. Buishand and Brandsma, 2001; Palutikof et al., 2002)) and one using regression. The limitations of these schemes is marked by imperfect representation of local daily variability in a future climate. Conducting a thorough comparison of the schemes with respect to hydrologic applications is overwhelmingly complex and can only be performed in small steps. Inspired by the last IPCC report (IPCC, 2001, 10.6.3.) the current study focuses on regression-based models. The report reflects on problems to represent point-wise daily variability and how to overcome them. Three approaches are cited that successfully restore the point-wise daily variability: randomization (RND, Buma and Dehn, 1998), inflation (IFN, Karl et al., 1990), and expanded downscaling (EDS, Bürger, 1996), all built on the regression scheme. There are two questions that we want to answer: First, to what extent does the restoration distort spatial correlations?—And second, how should one deal with the unexplained part in view of climate change?

For an observational network of temperature and precipitation stations from two river basins in Germany we simulate corresponding downscaled series, driven by observed and simulated atmospheric fields of the global atmosphere. In a final step, the simulated series of each basin drive a hydrologic model and generate the appropriate runoff. Our analysis is twofold: First, we compare various precipitation statistics and interpret them from their basic mathematical assumptions. Second, we try to understand the simulated runoff from these statistics and discuss additional effects, such as those stemming from temperature and temporal characteristics.

Only after this paper's first review a recent study (Wilby et al., 2003) came to our attention that has a similar intention by conducting a model intercomparison for multisite downscaling. The main differences are that, first, it has a focus solely on current climate and, second, the models are distinguished by different temporal behavior of one 'marker' site, from which the other sites are simulated using a resampling scheme.

2. Simulating local variability

All models under investigation are based on the regression scheme. To fully understand the generation of local variability we will describe that scheme with some detail, in particular its failure for hydrologic applications. Then follow the model recipes to reestablish local variability.

Each model defines a statistical relation between the large-scale, atmospheric process, g, and some local record of meteorological variables, ℓ. For the standard case of linear regression the assumed relation is of the form

$$\ell = \mathbf{L}g \tag{1}$$

with a matrix $\mathbf{L}$ that by definition minimizes the resulting (squared) model error. Specifically, if we denote the cross covariance of any two (multivariate) processes ξ and η by $\mathbf{C}_{\xi\eta}$, then the solution matrix $\mathbf{L_R}$ is given by the well-known linear regression term

$$\mathbf{L_R} = \mathbf{C}_{\ell g}\mathbf{C}_{gg}^{-1} \tag{2}$$

See Appendix A for a derivation of Eq. (2); there we also show how, in the presence of small global-local correlations, the model generates small amplitudes for avoiding large errors. For daily precipitation this correlation is particularly poor, rendering the model useless for downscaling applications.

2.1. How variability is added

We describe three approaches to deal with the unexplained portion of local variability. We emphasize that 'unexplained' here means 'linearly unexplained', but the exact definition of that portion will be given later in Section 3. Each approach utilizes some form of normalization. That means a basic state is estimated from observations and the model simulates anomalies relative to this state. Furthermore, since the large-scale atmospheric forcing fields are predominantly Gaussian, all precipitation data must be transformed to a (more or less) normal variate prior to any simulation. In this study, we use a generic normalization scheme for all models as described in Bürger (2002), cf. also Ledermann, (1984). It defines a 1–1 correspondence between daily precipitation series and normal variates with zero mean and unit

variance, using appropriate thresholds for dry days (so that the occurrence process is treated internally).

2.1.1. Randomization
The randomization model (RND) is an additive model of the form

$$\ell = \mathbf{L_R}g + \varepsilon \tag{3}$$

The error model, i.e. the simulation of an adequate process ε, can be defined in various ways. For example, it can be a simple white noise process of the Richardson type (Richardson, 1981) or a multivariate autoregressive model of order one (AR(1), cf. Brockwell and Davis, (1991). We applied the latter scheme; it is capable of memorizing local effects of the past day, but not more, and has correct spatial covariance.

2.1.2. Inflation
Instead of adding independent noise, the inflation model (IFN) alters the regression matrix itself, so that

$$\ell = \mathbf{L_I}g \tag{4}$$

gives correct local variance. This can easily be achieved by multiplying the jth row of $\mathbf{L_R}$ by $\sigma_j/\hat{\sigma}_j$, the ratio of observed and simulated standard deviation from the regression model. Note that by changing the ℓ-components in this way the mutual correlations between them are changed as well. It is interesting to note that IFN generates a smaller error than RND. A proof for the 1-dimensional case is given in Appendix B.

2.1.3. Expansion
EDS is an extension of the inflation model, having the same form as (4) but using a different matrix. Instead of a posteriori 'repairing' the regression model with appropriate scaling coefficients, EDS modifies the basic regression equation itself, in the following way: While for regression the matrix $\mathbf{L}$ should unconditionally minimize the (squared) model error, for EDS one introduces a side condition expressing the fact that the observed covariance be preserved (see Appendix C). This defines a nonlinear optimization problem that has a unique solution $\mathbf{L_E}$, called the EDS model. When driving EDS with

the global fields that were used for calibration (analyses), the entire covariance structure of the local fields is reproduced. This applies likewise to any atmospheric process g of the same climate (with respect to $\mathbf{C}_{gg}$). Such g generates realistic local variability that is representative of the present climate, at least as can be measured by the covariance matrix. That behavior is altered if the global covariance $\mathbf{C}_{gg}$ changes, for example in a GCM scenario. For details please consult Bürger, (1996).

Note that the RND model *implicitly disregards* the unexplained portion of atmospheric variability, which amounts to assuming a *constant* $\mathbf{C}_{gg}$! Climate change that unfolds in $\mathbf{C}_{gg}$ (i.e. in variability) can therefore not be covered by this model. Moreover, being purely stochastic the memory of the model (with respect to short-term fluctuations) has faded after only one day. For hydrologic applications, as we will see below, this can have major implications.

2.2. The hydrologic model

For the continuous simulation of river discharge we have applied the 'Nordic' version of the conceptual rainfall-runoff model HBV (Sælthun, 1996), which represents a synthesis of various versions of HBV used in Scandinavia. In these countries, HBV has been in operational use for over 20 years. Moreover, successful applications of HBV are reported from some 30 countries with different climatic conditions (Bergström, 1992; Bergström, 1995). The main reasons to apply HBV for our study were the sound physical description of the main runoff generating processes without exceeding a certain level of complexity, and the relatively low computational resources required for a climate impact simulation run over 100 years or more. The HBV model has been thoroughly validated in connection with climate downscaling studies, using observed as well as simulated global climate (cf. Bronstert et al., 2002; Bürger, 2002; Menzel and Bürger, 2002).

The general structure of HBV consists of three model components: (1) Snow accumulation and snow melt, (2) the simulation of soil moisture and runoff, and (3) a response and river routing procedure. The model expects precipitation and temperature data on input. Usually, HBV is applied at daily time steps, but appropriate data availability allows a higher temporal

resolution. HBV is a lumped model, i.e. it describes the catchment as a spatially homogeneous system. Even though a catchment may be subdivided into 10 equal-area elevation zones the parameters used are generally the same for all the sub-areas. Therefore, we decided to use the semi-distributed HBV-D, a derivative of the 'Nordic' HBV-model (Krysanova et al., 1999). This model version allows to classify a catchment into an optional number of sub-basins which are regarded as primary hydrological units. In analogy to the original model version the sub-basins are then classified into 10 height intervals, but each considering up to 15 vegetation types with a proper parameterization. For the two catchments of the current study, parameter calibration and further details are described in Schwandt (2004).

3. Downscaled precipitation

We compare the performance of the three downscaling models using observed and simulated atmospheric conditions. The observed fields were taken from the reanalysis project conducted by the National Center of Environmental Prediction (NCEP, cf. Kalnay et al., 1996) of the USA; we selected a grid defined by the rectangle between (10 W, 40 N) and (20 E, 60 N), for the following group of predictors variables:

(A) 500 hPa geopotential height
(B) 850 hPa temperature
(C) 700 hPa specific humidity

These variables reflect the main global agents that affect local precipitation: circulation (A), temperature (B), and moisture (C). The original NCEP grid size of 2.5×2.5 had to be interpolated to the slightly coarser GCM resolution of 2.8×2.8. All fields were projected onto the major principal components so that they hold 99% of the variance. The components (142 as a whole) were scaled to ensure that each field contributes equally to the final predictor variance. All fields were normalized relative to the annual cycle by centralizing and scaling with the respective means and (inverse) standard deviations.

The simulated fields were generated by the coupled atmosphere ocean GCM ECHAM4/OPYC3

of the DKRZ, using a T42 resolution (see Roeckner et al., 1996). From the model fields we selected the same area and variables as before for NCEP, and projected them onto the leading set of NCEP EOFs. By filtering out the seasonal cycle we obtained the final predictor set for the subsequent downscaling. The filtering included a seasonal rescaling due to incorrect seasonal amplitudes in the GCM fields. Finally, therefore, the first two moments of the GCM predictor fields (of some undisturbed 30-year base period) were made identical to the analyses; higher moments were not considered.

There are two simulations of the GCM: One with stationary climatic conditions (300 year 'control run') and one driven by the IPCC emission scenario IS95a, usually termed 'business as usual'. This run starts in 1860 and continues to the year 2100. Before 1990, emissions are estimated from historic measurements while after they are taken from IS95a. Aerosol effects are not included here. For the scenario, the seasonal cycle is determined from the model period 1961–1990 while an arbitrary 30-year period is chosen for the control run. The downscaled results of the NCEP analyses, the control and the business-as-usual run of the GCM will be denoted by ANA, CTL, and BAU, respectively.

The analysis has been conducted for the Ruhr and the Neckar basin in Germany. The mostly mountainous Ruhr basin is located in the state of North-Rhine Westphalia east of the city of Cologne. In the predominantly maritime climate, precipitation over the basin mostly occurs through westerly cyclonic activity. From that basin we selected five climate stations recording temperature, together with another 53 precipitation gauges. The Neckar basin is located in the South-West of Germany with lesser maritime and westerly cyclonic influence. From that basin we selected six climate stations recording minimum, maximum, and average temperature, and further 37 precipitation records; all data were measured on a daily basis. Except for the validation experiment described in the next section, the models were calibrated using global and local data from the WMO base period of 1961–1990. This applies to

(A) the calculation of the annual cycle and its anomalies;
(B) the subsequent EOF calculation;

(C) the normalization step of the local variables;
(D) the definition of the matrices L_R, L_I, and L_E.

Any simulated climate change must be regarded relative to this base climate. The important step of normalizing the local variables was done using the *probit* approach (Ledermann, 1984). With the probit function one is able to transform any quantity to a normal variate with zero mean and unit variance. (The nonlinearity of the probit measures the deviation from normality, which for precipitation is obviously quite strong.) The probit parameters define the climatology of that quantity. To avoid artificial expansion of low frequencies, such as slow climatic trends, by the models IFN and EDS, all datasets have been high-pass filtered using a threshold of one month. Only the shorter time scales are modeled after Section 2.1; the longer time scales are simulated using linear regression. Being regression-based, the definition of any of the three models requires no more than the knowledge of the first two moments of the global and local variables. This renders them quite robust against overfitting. For independent validation of the EDS model see below, and more thoroughly in (Bürger, 2002); this source also provides further details about the method.

3.1. Current climate

In Fig. 1 we see a typical winter and summer simulation of areal precipitation, P, for the Ruhr basin; for this case, the model was calibrated with independent data. Especially for summer it is evident that fast observed fluctuations are not reproduced by RND for which those are stochastic; slower variations (month to month and longer) are better reproduced. The performance of IFN and EDS is comparable in many respects, with IFN often being stronger than EDS, especially in winter. The following two figures demonstrate that this is a systematic behavior. Fig. 2 depicts the daily areal average of the full Ruhr dataset in the form of a qq-plot, distinguished by the winter (Oct–Mar) and summer (Apr–Sep) season. To account for its stochastic nature we actually used 10 realizations of RND. One sees that for winter the spread of RND rather closely follows the diagonal of perfectly reproduced scales. This is similar in summer, but with a larger spread and a tendency to
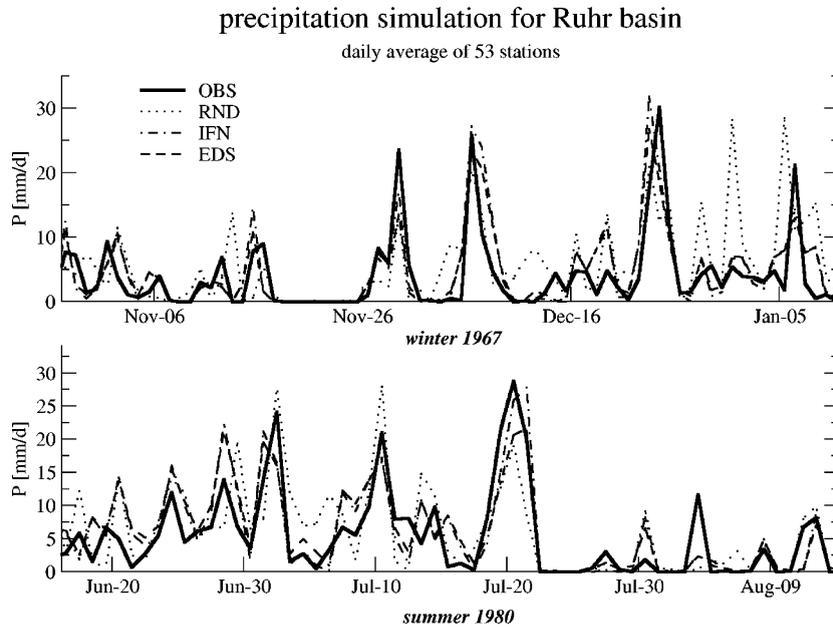
Fig. 1. Precipitation simulation for the Ruhr basin, for typical time spans. For winter 1967 (upper panel) it is evident that fast fluctuations of RND (dots, they are stochastic) are uncorrelated to the observations (OBS, heavy line), whereas larger clusters are better reproduced. IFN (dashed-dots) and EDS (dashed) perform satisfactory on both temporal scales, but in most cases IFN is bigger. The summer 1980 simulation (lower panel) appears more erratic. The heavy event around July 20 is reproduced quite well by IFN, and too weak by EDS and especially RND. Precipitation is shown as the average over.
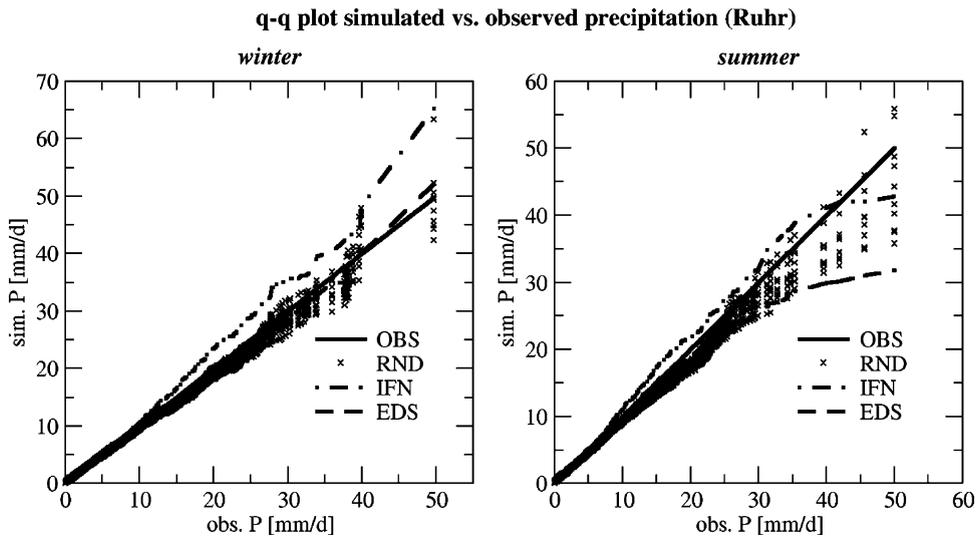
Fig. 2. qq-plot of daily observed vs. simulated precipitation P, areally averaged for the Ruhr basin. RND (crosses) comprises 10 realizations. For winter (left panel), both RND and EDS (dashed) reproduce observed (OBS, heavy line) scales with good accuracy, while IFN-simulated values (dashed-dots) are too large. For summer, the situation is somewhat reversed, with more accurate levels on the side of IFN and too low values, especially for scales greater than about 25 mm/d, for EDS. RND performs best on all scales.

q-q plot simulated vs. observed precipitation (Neckar)
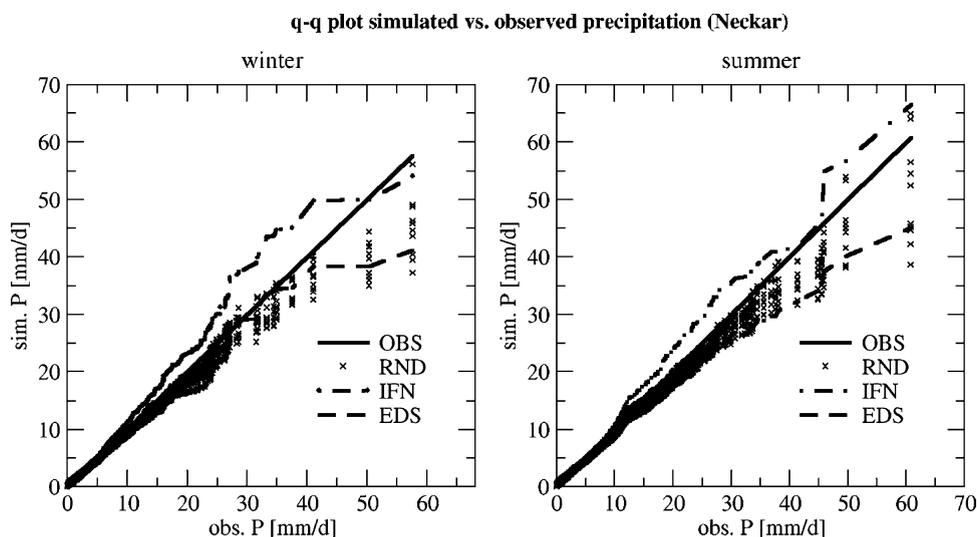
Fig. 3. Like Fig. 2, for the Neckar. The overestimation of IFN scales is apparent for both seasons. Winterly EDS scales are satisfactory except the extremes which are too weak, similarly to RND. Summerly RND and EDS scales are too weak either.

underestimate the strong events. Note that the RND spread is generated by some fixed statistical procedure with parameters that are calibrated to match the natural variation. Because systematic RND errors would show up as a bias the spread approximates the uncertainty that stems from natural fluctuations and/or estimation errors. That means, a very rough measure for the confidence strip about the diagonal of the qq-plot (i.e. observations) is given by the width of the RND spread of the corresponding observed scale. This uncertainty, at least, must be taken into account when assessing the performance of RND and the other models. Moreover, while 'true' distributions are smooth our shown discrete estimates are spoiled by sampling errors. This applies in particular to the extremes, including those from observations (see below).

On this background, the IFN model strongly overestimates all winterly and many of the summerly scales, while the EDS is almost perfect in winter but too weak in summer. For the Neckar (Fig. 3), the overestimation of IFN scales is even stronger. For winter, RND and EDS show similarly good performance, with exceptions for the very extremes, which are too weak. Summerly medium to large scales are too weak for both models. The extremes are more uncertain compared to the Ruhr. Note that the rightmost vertical RND column represents

the maximum observed P; especially for the Ruhr winter and Neckar summer this might well be an outlier (a sampling effect) so that the curves are somewhat misleading. Since the IFN is designed to preserve the local scale of a single variable, its obvious failure must be attributed to forming the areal average. In Appendix D we prove that the IFN model does in fact not preserve the scale of areal means, due to a misrepresentation of spatial correlations. Whether they are higher or lower than observations depends on the context. In our case they are too strong, as Table 1

Table 1

Average inter-station correlation, in %, in exceeding the 50- and 99%-percentile (that is, Kendalls τ_b), for observed and simulated precipitation, based on 1961–1990

		OBS	RND	IFN	EDS
Ruhr					
Winter	Q50	75	55	81	59
	Q99	49	33	69	37
Summer	Q50	73	53	79	53
	Q99	35	30	54	21
Neckar					
Winter	Q50	75	51	80	53
	Q99	49	30	64	19
Summer	Q50	73	51	79	50
	Q99	31	25	56	23

In all cases, RND (one realization) and EDS are too low and IFN is too high. The correlations are better reproduced for the Ruhr than for the Neckar.

shows. Using for any station as a binary measure the exceedance of the 50% (Q50) and the 99% (Q99) percentile, respectively, the Table shows their mutual correlation (or equivalently, Kendalls τ_b) averaged over all station pairs. In contrast, the RND and EDS simulated correlation is too weak, corresponding to the summerly averages from Figs. 2 and 3. Note in particular the very low Q99 values of EDS for the Neckar winter, which evidently conflict with the claimed preservation of local covariance by EDS.

Since the correct working of the EDS core module (Eq. (C1)) is without doubt this failure must be attributed to deficiencies in the normalization step. We suspect that the driving NCEP fields are not fully Gaussian—they were only normalized via mean and standard deviation (and not via the probit)—so that non-Gaussian distribution tails are causing these effects. This can of course be remedied, but it is certainly a drawback of the EDS model to be so sensitive to the details of that distribution.
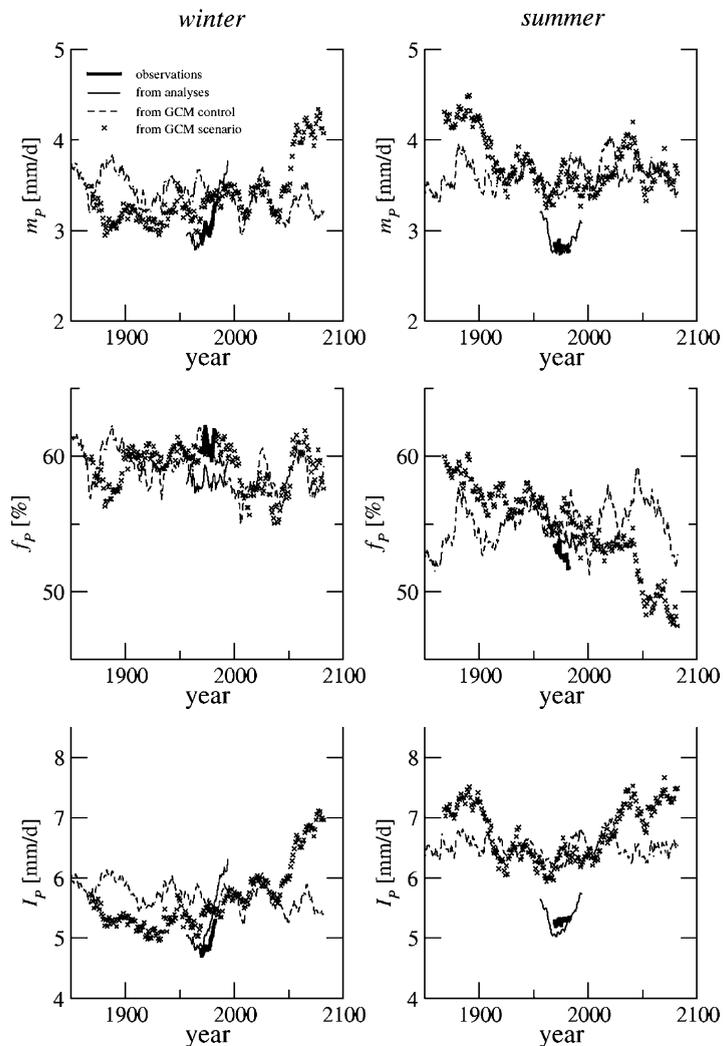
Fig. 4. Precipitation characteristics for various EDS simulations for the Neckar. The present winter (heavy line) climate is well reproduced. Summerly (right panel) m_P is only correct for ANA (thin line) but too strong when forced with the GCM (dashed: CTL; crosses: BAU); f_P is slightly too high in all models and I_P too low for ANA. For the BAU scenario, winterly increase is obvious for m_P and I_P. The summer picture is twofold, as there is a positive signal only for I_P accompanied by a negative signal in f_P.

3.2. Future climate

The former section defined the internal error of the various downscaling methods by using 'perfect' large-scale forcing fields. This error will be superimposed by external errors in the GCM-simulated large-scale forcing. To detect and analyze climate change signals more safely it is useful to have at least one further independent simulation of the GCM, such as the control run. Systematic errors are identical in both and should become visible in the downscaling result.

With respect to the effects of global warming on precipitation we focus on the two parameters of daily frequency and intensity. If there are such effects they are likely to have different characteristics in winter and summer. In this section, we concentrate on the comparison of RND and EDS; the IFN simulations are characterized by overestimated amplitudes known from the former section, while they are otherwise similar to EDS.

The main results of the EDS simulation are shown in Fig. 4 for the Neckar basin; the Ruhr results are comparable. The Figure depicts, for the winter and summer season separately, smoothed annual averages of daily mean (m_P), frequency (f_P) and intensity (I_P) of precipitation, formed as an average over all precipitation stations. Several things are noticeable.

- ANA-simulated I_P is somewhat weak.
- Present summer values of m_P and I_P are too strong for the GCM driven simulations CTL and BAU, the latter between 1961 and 1990; this can be attributed to errors in (the covariance structure of) the simulated summerly atmosphere.
- The CTL simulation is stationary (and does not differ markedly between the three models).
- Winterly m_P and I_P are projected to increase in the BAU scenario while f_P remains stationary.
- The positive signal for summerly I_P is balanced by a negative signal of summerly f_P, leading to a zero net effect for m_P.

To focus on the consequences of global warming we now show the same statistics for the BAU scenario only, as simulated by all models RND, IFN, and EDS. For RND we again used 10 realizations. As before, the Neckar scenarios are sufficiently representative of both basins, so we only show those in Fig. 5. All models reproduce the present climate statistics quite satisfactory, except for summerly I_P which is too strong (cf. Fig. 4). For the future, all models equally project the winterly statistics with stationary frequencies and increasing intensities, along with a decrease of summerly frequencies. They fail to consistently project the evolution of summerly intensities: While RND-simulated intensities remain stationary throughout the model period EDS simulates increasing I_P, starting with the 21st century. This raises the following questions:

(A) Can we understand the behavior of f_P and I_P?
(B) Why does EDS simulate summer intensification but not RND?

ad (A): Signals in f_P are most likely caused by changes in the large-scale atmospheric circulation, such as the occurrence and duration of cyclonic flow or blocking. Therefore, slow variations in the forcing fields are the main drivers for f_P characteristics. Slow variations are handled equally in all models, see above, which therefore produce similar f_P signals. The intensification of single precipitation events is likely caused by the enhanced water holding capacity of warmer air (cf. Trenberth, 1999) so that more water can precipitate per event. This effect is multiplied by another effect of global warming: enhanced evaporation and advection of moisture from the global oceans will increase the supply of water as well.

ad (B): Only the (linearly unresolvable) short-term variability of the atmosphere is treated differently in RND and EDS, and can thus be accounted for their deviating results: While that variability is ignored by RND and replaced by short-term noise of *constant* characteristics, the EDS model allows for the slow variation of that characteristic. This is possible as the EDS defines a *function* between the *statistics* of g ($\mathbf{C}_{gg}$) and ℓ ($\mathbf{C}_{\ell\ell}$), see Eq. (A4). Note that this function must not be mistaken, as von Storch (1999) does, as the model function between g and ℓ itself. One such variability signal is found by inspecting mean and variance of atmospheric humidity in the BAU scenario. Fig. 6 depicts the annual mean and standard deviation, calculated as the daily variation about the respective annual mean, of the leading principal
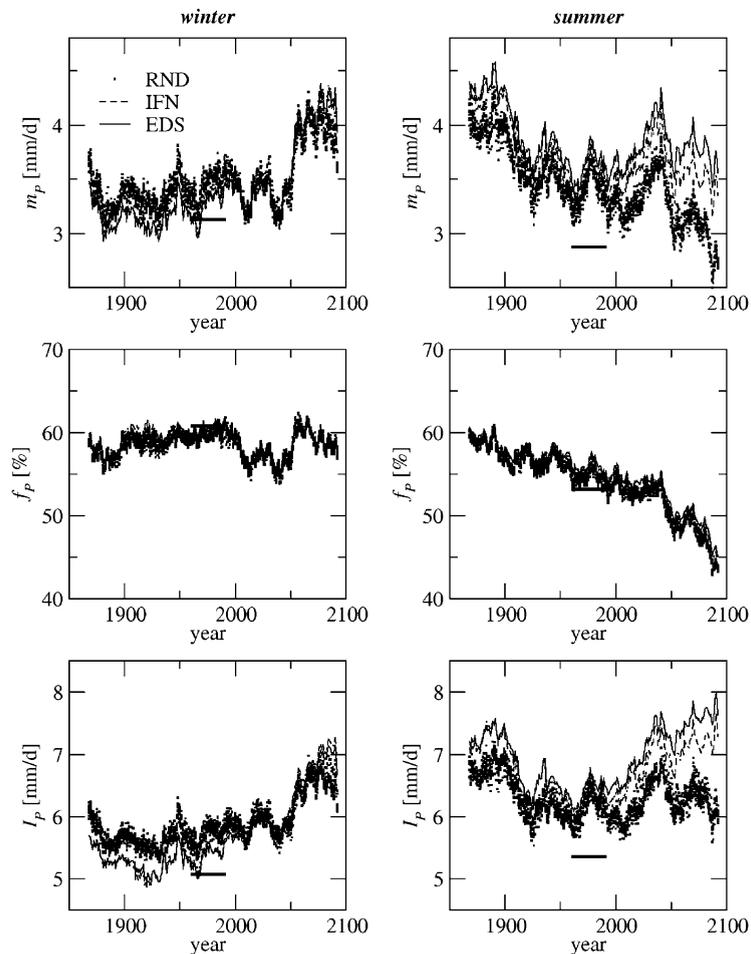
Fig. 5. Similar to Fig. 4, but only the BAU scenario by models RND (crosses), IFN (dashed), and EDS (solid). Present climate (black bars) is reproduced satisfactorily, although some summerly values, especially IFN, appear slightly too high. The models agree in projecting stationary frequencies and increasing intensities for winter and decreasing frequencies for summer. RND simulated I_P is stationary while IFN and EDS project an intensification.

component of the North-Atlantic/European specific humidity at 700 hPa. This principal component series is a rather prominent predictor field. Not only do we see a strong increase in the mean itself but also an increase in the variation about that mean. In this case, the increase is governed by the exponential growth with temperature of the water holding capacity of the atmosphere (from the Clausius–Clapeyron equation), and so is a direct nonlinear effect. Another possible variability signal, and certainly not the last, is a shift in the storm tracks, as a change in variance of the geopotential height fields. All these signals would

have a strong effect on local precipitation which cannot be captured by the RND model.

The different projections for summer intensities become even more pronounced if one looks at extreme events. To conduct an extreme value analysis we have to observe the fact that each of the results is based on the WMO baseline period of 30 years used for calibration, so that robust estimates of return periods will be confined to that scale at most. We have conducted analyses for a number of 30-year time spans. To utilize the full spectrum of GCM variability the control run was split into 10 adjacent pieces,
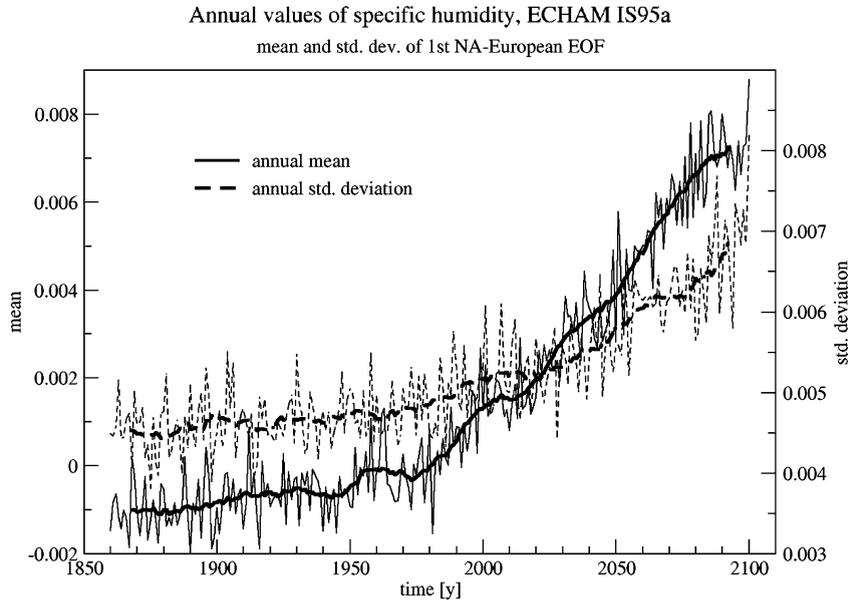
Fig. 6. BAU-scenario of atmospheric moisture, in terms of annual mean (solid) and standard deviation (dashed) of the first EOF of specific humidity (taking a North-Atlantic section from the ECHAM IS95a scenario run); heavy lines indicate smoothing.

which increases the sample size considerably. This results in 13 different 30-year spans, as follows:

- 1961–1990 observations (OBS)
- 1961–1990 simulations from atmospheric observations (ANA)
- Ten 30-year spans from the control simulation (10×CTL)
- 2061–2090 from the global warming scenario (BAU)

For each time span, we calculated the respective empirical cumulative distribution function (*cdf*) for the summer season. To acknowledge for the involved uncertainty we have not applied any function fits (such as Gumbel) to the groups of discrete points. Fig. 7 shows, exemplified for the Neckar basin (the Ruhr shows essentially the same), the resulting 13 *cdf*s for both models RND and EDS. Note the increase of sampling errors as one approaches the maximum of 30 years (for which the observed maximum is the only sample). Estimates of return periods below about 5 years are more robust. For both models we note a persistent underestimation of 1–5-year events in the ANA simulation; this corresponds to the underestimation of these scales seen in Figs. 3 and 4.

The 10 *cdf*s of the CTL run indicate an overestimation on a wide range of scales, in particular for summer. A closer investigation revealed as the main source of error incorrect correlations among the GCM predictor variables. Possible improvements are on the side of the GCM and hence beyond our control. Besides the sampling errors mentioned above, the extreme events (>5-year return period) show considerable spread in the 10 *cdf*s, indicating rather strong fluctuation. For RND, the BAU scenario is not distinguishable from that 'natural background'. EDS, however, simulates a clear intensification on all scales, with increases relative to the control run of up to 10 mm/d for a typical 1-year event. The following section will show if and how this affects the runoff generated from these scenarios.

4. Downscaling runoff

In the former section we have analyzed the performance of the three downscaling models with respect to current as well as future precipitation climate. As we saw, precipitation climate has many aspects, such as occurrence or intensity properties
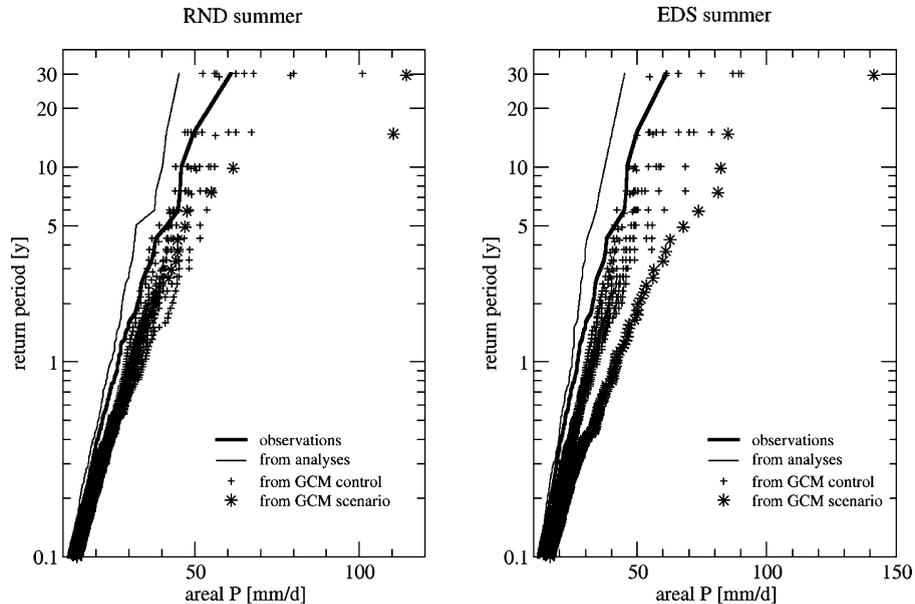
Fig. 7. Extreme value analysis of areal P for the Neckar, using 13 different cdfs, for summer simulations of RND (one realization, left panel) and EDS (right panel). Compared to observations (heavy solid line), events of return periods >1 year are too weak for ANA (thin line) in both models, while the 10 CTLs (crosses) are too strong. The BAU scenario (stars) is not significantly different for RND but shows a strong intensification for EDS.

along with spatial coherence, and we have tried to cover them as good as possible. We further saw that none of the models was perfect in all these aspects, especially when spatial properties were included. In this section, we feed the precipitation results into the hydrologic model and try to assess the importance of spatial precipitation coherence for its simulations. Note that to contain the scope of this paper temporal autocorrelation like, e.g. the duration of precipitation events was not at all considered.

To obtain credible results from the model chain GCM → downscaling → hydrologic model the following two things are crucial: First, present river climate should be reproducible from present atmospheric climate via downscaling and hydrologic modeling; second, simulated future trends of river climate should be as consistent as possible with a given global climate projection. In the above chain, the HBV is the least problematic. Previous studies (cf. Menzel and Bürger, 2002; Schwandt, 2004); see also http://dfnk. gfz-potsdam.de) demonstrate that the error produced by HBV is negligible against the errors in the input; hence we simply ignore them. Accordingly, by

referring to runoff 'observations' we generally mean HBV-simulations using observed precipitation.

4.1. Current climate

The following Fig. 8 corresponds with Fig. 1. For the same winter and summer periods we show the resulting Ruhr-runoff. The winterly overestimation by IFN is evident from the December 26 peak discharge. Here EDS appears slightly better than RND. The summer period of 1980 is best simulated by EDS, although the peak event of July 22 is not sufficiently pronounced. The respective Nash–Sutcliff coefficients, corresponding to the entire period 1961–1990, are for winter: IFN: 80%, EDS: 83%, and for summer: IFN: 79%, EDS: 82% (RND with 72 and 76%, respectively for one specific realization, is not comparable here due to its stochastic nature). The overall representation of scales is shown in Fig. 9, as an analogue of Fig. 2 with runoff replacing areal average precipitation. For winter, we note that the moderate scales are overestimated by IFN. For the extremes, i.e. scales beyond about 15 mm/d, the RND
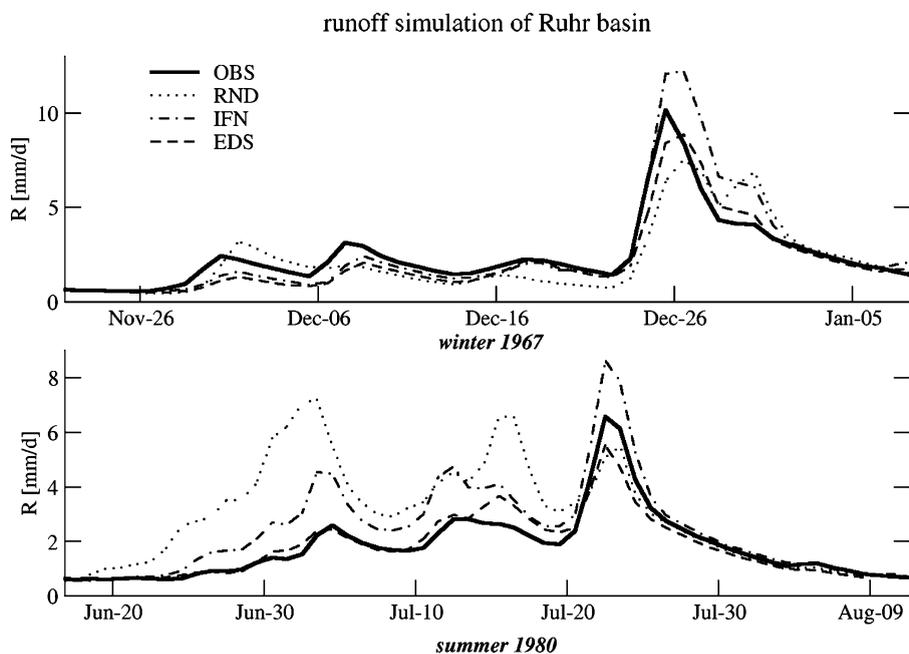
Fig. 8. This corresponds with Fig. 1, here with HBV generated runoff from the (spatially explicit) precipitation simulations of the three downscaling models. Relative to OBS (i.e. HBV-simulation from observed precipitation) the December 26 high flow (upper panel) is too strong in the IFN model and too weak especially for the RND model. The event around July 20, 1980, generated a high flow that is too strong in IFN, and slightly too weak in the other. Note the two strong 'artificial' events in earlier July.
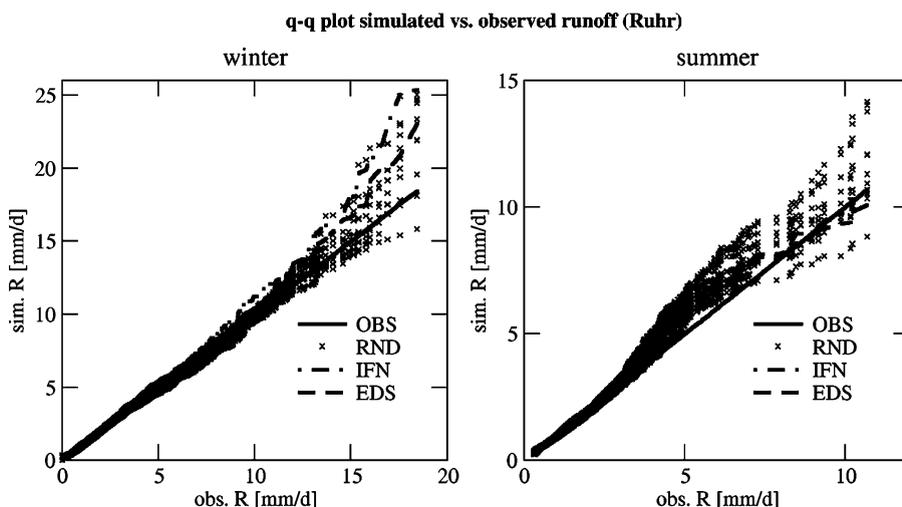
Fig. 9. Like Fig. 2, for the Ruhr runoff R. The winterly overestimation of R by IFN on all scales is obvious, and for flows larger than about 15 mm/d the same is true for the EDS. The RND simulations (crosses) do not exhibit any significant bias. For summer, all models overestimate moderate flows around 5 mm/d. Note that the summer statistic is much noisier.
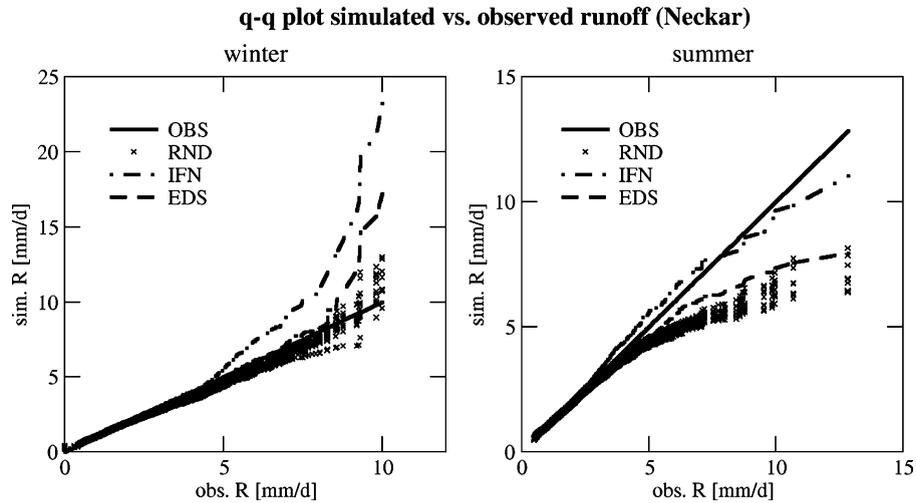
Fig. 10. Like Fig. 9, for the Neckar. Strong overestimation of extreme R-scales by IFN and EDS. But note that the four strongest events all belong to one simulated flood. For summer, the moderate events are too strong for IFN, and the extremes are much too weak for RND and EDS.

spread envelopes all other simulations, observations being roughly at the lower third and IFN at, and EDS slightly below, the top. Therefore, a significant departure from the present winter climate can at most be stated for the extremes of IFN and EDS. For example, four of the largest six simulated events (by IFN and EDS) belong to just one particular flood!— Many aspects can be understood from the areal average seen in Fig. 2, except the extremes of RND and EDS which appear too strong for runoff. The question of significance left aside—note that the runoff spread is considerable, their different behavior might result from discrepancies in the temporal domain (see below). Note again the overestimation of all scales in the IFN model. Summerly flows are similar in all models. It is interesting—but beyond our grasp—that they all show a slight positive bias for medium scales of about 5 mm/d. The uncertainty strip for the extremes (as indicated by the RND spread) is even larger here, and all simulations lie well within that strip. Comparing this to Fig. 2 it is noticeable that the underestimation of extremes by RND and EDS does not seem to have any negative influence on the runoff. This throws some light on the difficulty of the verification problem here: Namely, it is well possible that the negative spatial errors in those models are balanced by positive errors generated in the temporal domain. This is supported by the winterly figure

(including the Neckar, see below) where correct spatial statistics lead to an overestimation of runoff.

The results for the Neckar, shown in Fig. 10, are different. We note a striking overestimation of extreme winterly flow by IFN and EDS and a similar underestimation of summerly flow by RND and EDS. While for winter the 4 strongest simulated R-values belong to a single simulated flood, the too weak summerly RND/EDS flow is more systematic (cf. Table 1). Comparing this to Fig. 3 we find that it is the only case where a simulated precipitation scale is not 'amplified' by the hydrologic model (in the sense that too weak P generates correct R or correct P generates too strong R). Given that areal P is a good proxy for R one should, of course, try to explain the amplification rather than its absence. But as that is most likely related to deficiencies in the temporal domain of the regression method itself this attempt goes beyond the scope of the current study.

4.2. Future climate

The downscaled BAU climate scenarios as described in Section 3.2 were used as forcing of the HBV model. It turned out that the IFN- and EDS-driven projections were rather similar, so we treat them together. The resulting runoff R for the Ruhr basin is depicted in Fig. 11 as follows: For each
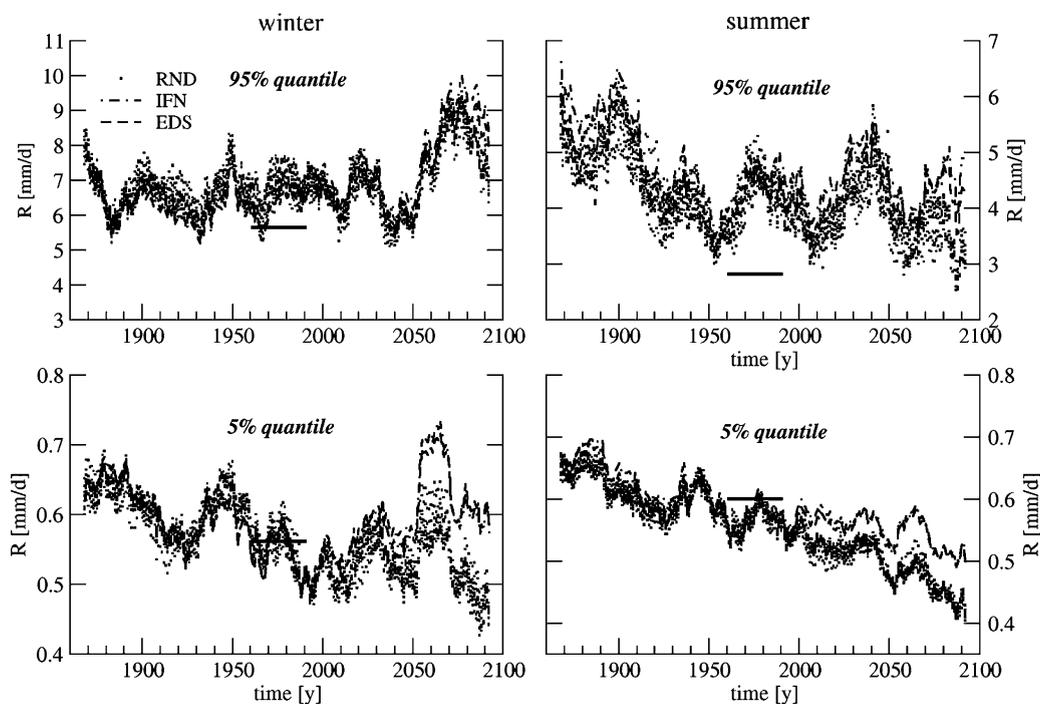
Fig. 11. For the Ruhr, low (5%-quantile, lower panels) and high (95%-quantile, upper panels) annual flow from the BAU scenario from the RND (crosses), IFN (dashed-dots) and EDS (dashed) model, in terms of runoff R (units: mm/d) using a 17-year moving average; left panel: winterly flow, right panel: summerly flow. Current flow climate (heavy solid line) is reproduced successfully by all models, except summerly high flow which is too weak in all models. Note that all models fluctuate remarkably synchronous. Winterly high flow is consistently projected to increase as well as decreasing summerly low flow. But starting at about the year 2000 IFN/EDS diverge from RND with larger (especially low) flows.

simulated year between 1860 and 2100 the low flow (5%-quantile) and the high flow (95%-quantile) are shown separately for the winter and summer season. The present flow climate is reproduced quite well in all models, except summer high flow which is slightly too large. All models agree qualitatively in projecting increasing winterly high flow and decreasing summerly low flow. The strength of the signals are, however, different. For example, the summerly decrease of low flow is much more severe for RND. Furthermore, RND-simulated summerly high flow at the year 2100 is about 1 mm/d smaller than that of IFN/EDS.

The differences are even more pronounced for the Neckar, shown in Fig. 12. In the IFN/EDS simulations for both seasons, the negative trend in low flow is turned positive after about 2000. The same applies to high flow, with runoff almost doubled in IFN/EDS compared to RND at the year 2100. The parallels to Fig. 5 are notorious, suggesting that one can attribute

the rise in summerly Neckar flow to a rise in precipitation intensities.

5. Conclusions

We have analyzed the simulations of 3 different downscaling models, RND (randomization), IFN (inflation), and EDS, and compared their results using subsequent hydrologic simulations. As a global forcing we used observed as well as simulated atmospheric fields, the latter representing either the present or a warming climate. Not surprisingly, the results are quite diverse and not easy to stratify. A great complication arises from the superposition of spatial and temporal correlation errors in the downscaled fields, and the possibility that both cancel each other in the runoff simulation. Runoff alone is then a bad indicator in the assessment of climate scenarios. For successful simulations of the present climate
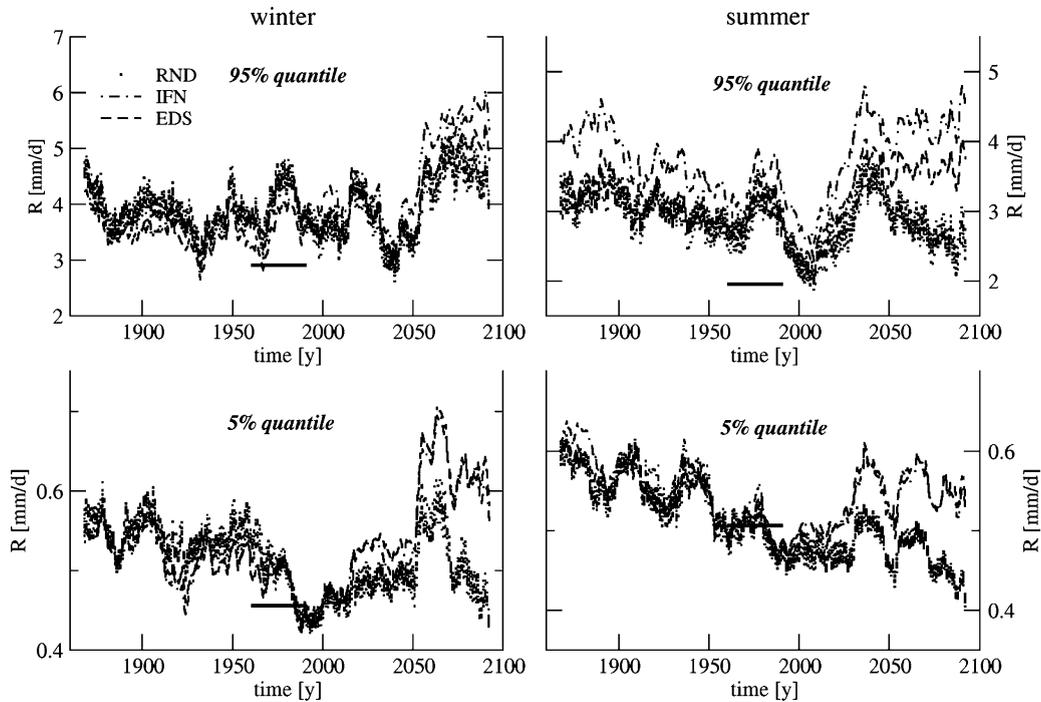
Fig. 12. Like Fig. 11, for the Neckar. Present runoff climate is within the standard error (not shown), except for summer high flow which is too strong. Winterly high flow is projected to increase in all models, while RND simulates stationary and IFN/EDS increasing low flow. A decrease of summerly low flow is only simulated by RND whereas under IFN/EDS it remains stationary. High flow increases for IFN/EDS and remains stationary for RND.

might be stemming from incorrect precipitation statistics, whose errors only affect scenarios of future climate. Our study focused on the analysis of the spatial properties of the three models and compared their results with the final runoff. It should be complemented by further studies with emphasis on the temporal behavior of these models.

None of the models was perfect even with respect to the spatial behavior. We found it important to not only state the failures but also try to understand them. To summarize, we found that the RND model showed very good performance with respect to present climate, but was inherently unable to reflect changes in the atmospheric variability. And we presented evidence that such changes are to be expected in a warming climate. The IFN model is the most simple of the three models. But for multisite downscaling it is flawed as it misrepresents spatial correlations. This leads, for example, to incorrect scales in the simulated areal precipitation. The EDS model might in theory avoid the errors of the other two, but it turned out that

its performance heavily depends on the details of the normalization. Because the driving atmospheric fields were not perfectly transformed to Gaussian variates the model produced considerable errors in the spatial correlations. On the other hand, of all models it might suffer most from biases in the simulated driving fields. The model will certainly profit from an improvement in the normalization.

The simulated runoff revealed other deficiencies whose origin is partly unclear: Winterly high flow climate is too strong for IFN and EDS, compared to observations, although this is the result of only one simulated flood in the case of the Neckar. And summerly high Neckar flow is much too weak in RND and EDS. These failures are likely to be related to the overall climatic conditions of the respective basins. Since the Ruhr has a closer proximity to the North Sea its climate is more maritime and strongly determined by the circulation over the Atlantic. For the mountainous catchment of the Neckar the small-scale local effects are more relevant which are harder to relate to

(downscale from) that circulation. Therefore, differences and errors in the mechanism to generate the residual portion of local variability will show up more pronounced for this basin. We can only speculate as to what causes these deficits. For example, the temporal structure is misrepresented that produces the typical sequence of events prior to a big flood. Or, the local normalization scheme (cf. Section 2.1) that transforms the extreme values is too badly parameterized so that small errors are systematically accumulated. Furthermore, as indicated above, all models might negatively be affected by the residual non-normality of the atmospheric fields.

Despite the deficiencies in reproducing the present climate the global warming projections are still remarkably consistent, except for the above-mentioned case for RND. All models project an intensification of winterly precipitation with stationary frequencies, leading to increased winterly high flow when driving the hydrologic model. They also project a rather strong decrease in the frequency of summerly rainfall, which in turn might be attributable for the projected negative trend in summerly low flow. Evaporative effects through enhanced warming do certainly play a major role here, too. For high flow, rainfall intensity is more relevant, but here the models diverge: While remaining stationary under RND the IFN/EDS-simulated intensities show a marked increase. We have attributed this increase to a strong rise in the variability of atmospheric moisture as simulated by the GCM. As described above, this signal cannot be captured by RND.

Present summerly high flow was severely too weak for the RND and EDS model and, as we suspect, correct for IFN only because of a positive intensity-bias. This points to a general deficiency in the regression approach. To investigate this further a comparison with other, for example weather-type or resampling schemes, is advisable. One must always be aware, however, that the calibration of any downscaling model is limited to periods of no more than a few decades, as given by the availability of atmospheric observations (analyses). Therefore, robust statistics such as return periods of river floods, for example, can only be derived for that order of magnitude, and everything beyond is too uncertain. For example, the 10 RND realizations as well as the 10 CTL simulations showed clearly that even

the 30-year return period of the maximum flow is loaded with uncertainty.

Appendix A. Reduction of variance

Each model defines a statistical relation between the large-scale, atmospheric process, g, and some record of hydrologic variables, ℓ. For the standard case of linear regression the relation is of the form

$$\ell = \mathbf{L}g \tag{A1}$$

with a matrix $\mathbf{L}$ that is determined to minimize the resulting model error. Specifically, if we denote the cross covariance of any two (multivariate) process ξ and η by $C_{\xi\eta}$, then $\mathbf{L}$ is defined as the matrix that minimizes the trace of the error covariance $C_{\varepsilon\varepsilon}$, with $\varepsilon = \mathbf{L}g - \ell$. This is easily calculated as

$$\min_{\mathbf{L}}\{\mathbf{C}_{\varepsilon\varepsilon}\} = \min_{\mathbf{L}}\{\mathbf{L}\mathbf{C}_{gg}\mathbf{L}^T - 2\mathbf{L}\mathbf{C}_{gf} + \mathbf{C}_{\ell\ell}\} \tag{A2}$$

and solved by equating the resulting Jacobian to zero, which yields the well-known linear regression term

$$\mathbf{L_R} = \mathbf{C}_{\ell g}\mathbf{C}_{gg}^{-1} \tag{A3}$$

The local variance reduction of this model can be deduced as follows. First, we note that for any linear model $\mathbf{L}$

$$simulated\ local\ covariance = \mathbf{L}\mathbf{C}_{gg}\mathbf{L}^{\mathbf{T}} \tag{A4}$$

From the canonical correlation matrix

$$\mathbf{K} = \mathbf{C}_{\ell g}\mathbf{C}_{gg}^{-1}\mathbf{C}_{gf}\mathbf{C}_{\ell\ell}^{-1} \tag{A5}$$

one knows that its norm obeys $|\mathbf{K}| < 1$ (all eigenvalues are <1; cf. von Storch and Navarra, 1993). Hence, for the simulated local covariance of the regression model $\mathbf{L_R}$ it follows from (3) that

$$|\mathbf{L_R}\mathbf{C}_{gg}\mathbf{L_R^T}| = |\mathbf{K}\mathbf{C}_{\ell\ell}| < |\mathbf{C}_{\ell\ell}| \tag{A6}$$

In words: using linear regression, the scale of the simulated local variability (the left hand side of (4)) is bound by the prevailing global-local correlations, and smaller than the observed local variability (the right hand side of (4)).

Note that using (A2) and (A5), the error covariance of linear regression is

$$\mathbf{C}_{\varepsilon\varepsilon} = (1 - \mathbf{K})\mathbf{C}_{\ell\ell} \qquad (A7)$$

Appendix B. Model error

For 1-dimensional linear regression, (A7) becomes

$$\langle \varepsilon^2 \rangle = (1 - \rho^2)\sigma_\ell^2 \qquad (B1)$$

The inflation error variance is

$$\langle \varepsilon_{\mathbf{I}}^2 \rangle = \langle (\mathbf{L}_{\mathbf{I}} g - \ell)^2 \rangle = 2\sigma_\ell^2(1 - \rho) \qquad (B2)$$

Both are related by ρ as follows:

$$\frac{\langle \varepsilon_{\mathbf{I}}^2 \rangle}{\langle \varepsilon_0^2 \rangle} = \frac{2}{1 + \rho} \qquad (B3)$$

For ℓ = precipitation, ρ lies in the order of 0.5 (using a suitable g), so that roughly $\langle \varepsilon_{\mathbf{I}}^2 \rangle = 1.3\langle \varepsilon_0^2 \rangle$. The randomization model always gives $\langle \varepsilon_{\mathbf{R}}^2 \rangle = 2\langle \varepsilon_0^2 \rangle$, which is considerably larger than the inflation error.

Appendix C. Expanded downscaling

With the notation of Appendix A, the basic definition of EDS is as follows

$$\min_{\mathbf{L} \in S}\{\mathbf{L}\mathbf{C}_{gg}\mathbf{L}^{\mathbf{T}} - 2\mathbf{L}\mathbf{C}_{g\ell} + \mathbf{C}_{\ell\ell}\}$$
$$\qquad (C1)$$
$$S = \{\mathbf{L}|\mathbf{L}\mathbf{C}_{gg}\mathbf{L}^{\mathbf{T}} = \mathbf{C}_{\ell\ell}\}$$

which simply means that one seeks a matrix $\mathbf{L}$ that minimizes the model error as in regression, but subject to the constraint that the simulated covariance be preserved. Eq. (C1) describes a nonlinear constraint minimization problem (nonlinear programming problem). It can be shown that it has a unique solution $\mathbf{L_E}$ which we refer to as the EDS model. We note the following:

- Disregarding the second equation results in the standard linear regression model.

- Disregarding the off-diagonal elements in the second equation results in the inflation model.

The latter is not trivial (relative to the authors analytical skills): Why should the inflation model not be sub-optimal in the sense of (C1). But random tests showed it is in fact optimal.

Appendix D. Spatial correlations

For simplicity, we assume that we only have two local variables, ξ and η, with mean zero. The expectation of the sum is

$$\langle (x + y)^2 \rangle = \mathbf{C}_{\xi\xi} + 2\mathbf{C}_{\xi\eta} + \mathbf{C}_{\eta\eta} \qquad (D1)$$

The same equation applies to the simulated variables. The INF model preserves the two outer terms, $C_{\xi\xi}$ and $C_{\eta\eta}$, of the right-hand side of (D1), describing the variance of the variables ξ and η, respectively. The mixed terms, $C_{\xi\eta}$, are not preserved, nor is, hence, the overall expression. While for the univariate case it can be shown that the sum is generally expanded, the situation for multivariate cases seems more complex, so that the shrinking or expanding of the sum depends on the context.

References

Bardossy, A., Plate, E., 1991. Modeling daily rainfall using a semi-Marcov representation of circulation pattern occurrence. J. Hydrol. 122, 33–47.

Bardossy, A., Plate, E., 1992. Space-time model for daily rainfall using atmospheric circulation patterns. Water Resour. Res. 28, 1247–1259.

Bergström, S., 1992. The HBV model—its structure and applications, SMHI Report RH, Norrköping, No. 4.

Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Littleton, CO.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods, second ed. Springer, Berlin.

Bronstert, A., Bismuth, C., Menzel, L. (Eds.), 2000. European Conference on Advances in Flood Research Proceedings (vols. 1–2), Potsdam-Institute for Climate Impact Research, Potsdam, Germany. Summary Report 65.

Bronstert, A., Niehoff, D., Bürger, G., 2002. Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities. Hydrol. Process. 16, 509–529.

Buishand, T.A., Brandsma, T., 2001. Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. Water Resour. Res. 37 (11), 2761–2776.

Buma, J., Dehn, M., 1998. A method for predicting the impact of climate change on slope stability. Environ. Geol. 35, 190–196.

Bürger, G., 1996. Expanded downscaling for generating local weather scenarios. Clim. Res. 7, 111–128.

Bürger, G., 2002. Selected precipitation scenarios across Europe. J. Hydrol. 262, 99–110.

Charles, S.P., Bates, B.C., Hughes, J.P., 1999. A spatiotemporal model for downscaling precipitation occurrence and amounts. J. Geophys. Res. C4 104 (D24), 31657–31669.

Giorgi, F., Mearns, L.O., 1999. Introduction to special section: Regional climate modeling revisited. J. Geophys. Res.-Atmos. 104, 6335–6352.

Grabs, W., 1997. Impact of climate change on hydrological regimes and water resources management in the Rhine basin. CHR I-16, 172.

IPCC, 2001. Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Jacob, D., Andrae, U., Elgered, G., Fortelius, C., Graham, L.P., Jackson, S.D., Karstens, U., Koepken, Chr., Lindau, R., Podzun, R., Rockel, B., Rubel, F., Sass, H.B., Smith, R.N.D., Van den Hurk, B.J.J.M., Yang, X., 2001. A comprehensive model intercomparison study investigating the water budget during the BALTEX–PIDCAP period. Meteorol. Atmos. Phys. 77, 19–43.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetma, A., Reynolds, R., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-year reanalysis project. Bullet. Am. Meteorol. Soc. 77, 437–471.

Karl, T.R., Wang, W.C., Schlesinger, M.E., Knight, R.D., Portmann, D., 1990. A method of relating general circulation model simulated climate to observed local climate. Part I: seasonal statistics. J. Clim. 3, 1053–1079.

Kilsby, C.G., (Ed.), 2000. WRINCLE. Final report, Water Resource Systems Research Laboratory.

Kleinn, J., Frei, C., Gurtz, J., Vidale, P.L., Schär, C., 2002. Coupled climate-runoff simulations: A process study of current and a warmer climate in the Rhine basin. Institute for Climate Research, ETH Zürich.

Krysanova, V., Bronstert, A., Müller-Wohlfeil, D., 1999. Modelling river discharge for large drainage basins: from lumped to distributed approach. Hydrol. Sci. J. 44, 313–331.

Ledermann, W. (Ed.), 1984. Handbook of Applicable Mathematics A, vol. VI. Wiley, Chichester.

Machenhauer, B., Windelband, M., Botzet, M., Christensen, J.H., Deque, M., Jones, R., Ruti, P.M., Visconti, G., 1998. Validation and analysis of regional present-day climate and climate change simulations over Europe, Max-Planck-Institut fuer Meteorologie, Hamburg, Report 275.

Menzel, L., Bürger, G., 2002. Climate change scenarios and runoff response in the Mulde catchment (Southern Elbe, Germany). J. Hydrol. 267, 53–64.

Palutikof, J.P., Goodess, C.M., Watkins, S.J., Holt, T., 2002. Generating rainfall and temperature scenarios at multiple sites: examples from the Mediterranean. J. Clim. 15 (24), 3529–3548.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. 17, 182–190.

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., Schulzweida, U., 1996. The atmospheric general circulation model ECHAM4: Model description and simulation of present-day climate, Max Planck Institut für Meteorologie, Hamburg, Report 218.

Sælthun, N.R., 1996. The 'Nordic' HBV Model. Description and documentation of the model version developed for the project Climate Change and Energy Production, Norwegian Water Resources and Energy Administration. NVE Publication p. 7.

Schwandt, D., 2004. Abflußentwicklung in Teileinzugsgebieten des Rheins—Simulationen für den Ist-zustand und für Klimaszenarien, Potsdam-Institute for Climate Impact Research, Report 88.

Trenberth, K.E., 1999. Conceptual framework for changes of extremes of the hydrological cycle with climate change. Clim. Change 42, 327–339.

von Storch, H., 1999. On the use of 'inflation' in downscaling. J. Clim. 12, 3505–3506.

von Storch, H., Navarra, A. (Eds.), 1993. Analysis of climate variability Application of Statistical Techniques. Springer, Berlin, p. 6.

Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Prog. Phys. Geogr. 21, 530–548.

Wilby, R.L., Tomlinson, O.L., Dawson, C.W., 2003. Multi-site simulation of precipitation by conditional resampling. Clim. Res. 23, 183–194.

Wilks, D.S., 1999. Multisite downscaling of daily precipitation with a stochastic weather generator. Clim. Res. 11, 125–136.