Seasonal Predictability of European Discharge: NAO and Hydrological Response Time

M. F. P. Bierkens and L. P. H. van Beek

Department of Physical Geography, Utrecht University, Utrecht, Netherlands

(Manuscript received 1 April 2008, in final form 10 January 2009)

ABSTRACT

In this paper the skill of seasonal prediction of river discharge and how this skill varies between the branches of European rivers across Europe is assessed. A prediction system of seasonal (winter and summer) discharge is evaluated using 1) predictions of the average North Atlantic Oscillation (NAO) index for the coming winter based on May SST anomalies of the North Atlantic; 2) a global-scale hydrological model; and 3) 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) data. The skill of seasonal discharge predictions is investigated with a numerical experiment. Also Europe-wide patterns of predictive skill are related to the use of NAO-based seasonal weather prediction, the hydrological properties of the river basin, and a correct assessment of initial hydrological states. These patterns, which are also corroborated by observations, show that in many parts of Europe the skill of predicting winter discharge can, in theory, be quite large. However, this achieved skill mainly comes from knowing the correct initial conditions of the hydrological system (i.e., groundwater, surface water, soil water storage of the basin) rather than from the use of NAO-based seasonal weather prediction. These factors are equally important for predicting subsequent summer discharge.

1. Introduction

The ability to properly forecast river discharge at seasonal time scales is extremely beneficial to society. By searching for accurate predictions of river discharge for the coming season, one may improve reservoir management and help to ensure drinking water and food supply, hydropower generation, and river navigability (Trigo et al. 2004; Wilby et al. 2004; Cherry et al. 2005). This justifies the many attempts to relate river discharge to predictable, slowly varying components of the climate system, such as sea surface temperature (SST) and related large-scale atmospheric circulation patterns (Cullen et al. 2002; Peterson et al. 2002; Tootle et al. 2005).

In Europe, the North Atlantic Oscillation (NAO) is considered to be the strongest driver of interyear climate variability (Hurrell and van Loon 1997; Appenzeller et al. 1998) and the main factor explaining anomalies in winter precipitation and temperature (Hurrell 1995). Owing to the connection between Atlantic SST and the winter NAO and slow variation and inertia of the Atlantic Ocean, winter NAO can be estimated from May SST anomalies with reasonable skill (Rodwell et al. 1999; Rodwell and Folland 2002). Therefore, in regions with a strong relationship between the NAO and precipitation, seasonal prediction of river discharge may indeed be feasible. For these reasons, studies have focused on the relationship between the NAO and river runoff, including recent work on river flows in the British Isles (Wilby 2001), precipitation and discharge over the Iberian Peninsula (Trigo et al. 2004), discharge of the lower Danube (Mares et al. 2006) and discharge of Scandinavian rivers in relation to hydropower (Uvo and Berndtsson 2002).

Apart from predictability of the winter NAO and the related anomalies of winter precipitation and temperature, seasonal prediction of river discharge may be successful due to the slow response time of many large river basins. Rivers attached to large areas of lakes and wetlands (e.g., the Congo and Amazon) often have lag times between precipitation and runoff peak of weeks to even months (e.g., Dai and Trenberth 2002), whereas rivers that are strongly dominated by groundwater reservoirs (e.g., the Niger) may even show multiannual response times between wet periods and discharge (Zwarts et al. 2005). It can be expected that, if the initial
state (i.e., storage in lakes, rivers, groundwater, and soil) of these slowly reacting systems is estimated with sufficient accuracy, seasonal predictions may be attainable, even with average meteorological forcing.

The main goal of this paper is to assess the skill of seasonal prediction of European discharge and how this skill varies between the branches of European rivers. Also, we aim to assess how the skill depends on use of NAO-based seasonal weather prediction, the properties of the river basin, and a correct assessment of initial hydrological states. The NAO-based seasonal weather prediction, which will be described in detail in the next section, is a statistical approach in which historical meteorological fields (precipitation, evaporation, and temperature) based on 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) (Uppala et al. 2005) are related to a predicted winter NAO index based on May SST (Rodwell and Folland 2002). Seasonal forecasts of river discharge are obtained by running the predicted meteorological fields through a large-scale hydrological model (van Beek 2007). The method is first tested by performing a numerical experiment, and then we test the method through seasonal hindcasting of observed discharge data at 66 gauging stations across Europe.

2. Methods

a. Seasonal prediction method

Our goal is to predict river discharge of the coming season based on information from the preceding season. Our total season is a hydrological year with a winter season, [October–March (ONDJFM)] during which mean precipitation exceeds mean evaporation, and a summer season, [April–September (AMJJAS)] during which mean precipitation is generally less than evaporation. The forecasting method consists of three parts: 1) forecasting next winter’s NAO index; 2) selecting associated meteorological data; and 3) calculating next season’s (winter and summer) discharge.

Before we elaborate on these three steps, we will discuss the type of seasonal prediction method used. The method that we chose is a statistical prediction method that uses recorded years with similar large-scale and persistent atmospheric conditions (in our case the NAO phase) as analogs for the present. Quite a few studies showed large correlation between the NAO index and winter precipitation for different parts of Europe, notably in Scandinavia (correlation up to 0.8, Uvo and Berndtsson 2002) and the Iberian Peninsula (correlation up to −0.6, Trigo et al. 2004). This makes the NAO index a good candidate for seasonal prediction. We did not perform a comparison of this method with other more classical statistical methods (e.g., Wilby et al. 2004) and seasonal forecasts from dynamic approaches using coupled ocean–atmospheric general circulation models (OAGCMs). However, checking the reported skill scores of dynamic seasonal forecasts of winter rainfall over Europe at different prediction centers [e.g., ECMWF, National Weather Service (NWS)/National Centers for Environmental Prediction (NCEP), and Hadley Centre] for similar lead times (3–6 months) shows that results are not better than that of our simple method (average skill in terms of correlation is 0.2, from Fig. 5c in which we use runoff as a proxy for rainfall). Indeed, in one of the few reviews about approaches and skill of seasonal forecasting, Rodwell and Doblas-Reyes (2006) show that the current skill of multimodel seasonal forecasting of precipitation over Europe (3-month lead time) does not exceed a correlation of 0.2, even for the ensemble mean. So, although our conclusions about the skill of seasonal forecasts only pertain to the particular method used in this paper, we are confident that other seasonal prediction methods would have led to similar conclusions.

1) Forecasting the NAO Index

Rodwell and Folland (2002) devised a method to predict the strength of the DJF NAO index [the average pressure difference between the Azores high and the Icelandic low; we use the Hurrell (1995) definition] based on the North Atlantic SST anomaly observed in the previous month of May. The method is based on a lagged singular value decomposition of May North Atlantic SST and the subsequent average DJF 500-hPa geopotential height. This technique decomposes the two fields into orthogonal patterns such that cross-correlation between the patterns is at maximum. The first component of the average DJF 500-hPa geopotential height is similar to that associated with a positive NAO phase. By projecting an observed SST temperature anomaly on the first orthogonal predictor pattern of SST (multiplying the SST anomaly with the first SST EOF and summing the results), a single number results that is positive if a positive winter NAO index is predicted.

Figure 1 shows the forecasted DJF NAO index for the period 1957–2002 (the ERA-40 reanalysis period) taken from Rodwell and Folland (2002) together with the average winter (ONDJFM) NAO index from Hurrell (1995). It can be seen that the correspondence is reasonable, which is confirmed by the skill measured in terms of cross-correlation between the two curves: 0.55 for the entire ERA-40 period, 0.35 for the period 1961–80, 0.53 for the period 1981–2000, and 0.64 for the period 1991–2000. In our prediction method we will take the NAO index of Rodwell and Folland as given.
2) SELECTING ASSOCIATED METEOROLOGICAL DATA

The next step is based on the assumption that years with a similar winter NAO index will have similar precipitation, evaporation, and temperature winter anomalies. We take the ERA-40 daily reanalysis fields (Uppala et al. 2005) of precipitation ($P$), evaporation ($E$), and daily average temperature ($T$) as a historical archive and select for a given year the three ERA-40 hydrological years with an observed winter NAO index closest to the forecasted index for that year. These three years are then viewed as equally likely seasonal predictions of the $P$, $E$, and $T$ for the coming hydrological year. By taking three years instead of a single closest year, we reduce the effect of sampling error. In a short series of only 44 years, a few large errors in the predicted NAO index or associated discharge have a large negative impact on the accuracy; on the other hand, taking many years reduces the resolution (differences between years) too much. A sensitivity analysis of Scandinavian discharge resulted in similar results when taking two or four years. Based on this, we decided to take three years as a compromise between sampling accuracy and resolution.

3) CALCULATING DISCHARGE

To convert seasonal predictions of meteorological variables (from steps 1 and 2) to seasonal predictions of discharge, we need a hydrological model. Here we use the global hydrological model PCRaster Global Water Balance (PCR-GLOBWB) (van Beek 2007). PCRaster (Wesseling et al. 1996) is the scripting language in which the model is coded. The model simulates soil water storage, groundwater storage, and specific runoff (local runoff per unit land area) globally for daily time steps at $0.5^\circ \times 0.5^\circ$ spatial resolution. Specific runoff is accumulated along a drainage network derived from the 30' global drainage direction map (DDM30) dataset (Döll and Lehner 2002) that coincides with the major rivers of the world. The daily $P$, $E$, $T$ fields belonging to the three selected hydrological years from ERA-40 are used as input to PCR-GLOBWB to calculate river discharge for all stretches of European rivers. A model run starts at 1 October and is forced by the daily $P$, $E$, $T$ fields, which results in time series of daily discharge (accumulated runoff from upstream) for each river stretch. From this run the average winter (ONDJFM) and subsequent summer (AMJJAS) discharge is calculated for each river stretch and for each of the three selected years. Finally, winter and summer discharge for the three years is averaged, which results in the seasonal predictions of winter and summer discharge for each stretch of the larger rivers in Europe.

b. Numerical experiment

To assess the skill of the seasonal prediction method described above, we perform a numerical experiment with the ERA-40 data.

1) First, we simulate time series of daily discharge for the stretches of the European rivers by forcing PCR-GLOBWB with daily $P$, $E$ and $T$ fields from ERA-40 for the period 1957–2002. From this we assemble 44 years of winter and summer discharge for all river stretches for hydrological years 1957/58 to 2001/02. These are then considered the “true” discharge.

2) Next, for each hydrological year we use the predicted NAO index (see Fig. 1) to select three hydrological years from the ERA-40 time series with the NAO index closest to the predicted one and run these through the PCR-GLOBWB model to obtain time series of predicted winter and summer discharge for each river stretch. The skill of the method is then computed by the correlation coefficient between predicted and true discharge. This experiment is repeated twice:

- Using the exact initial conditions of PCR-GLOBWB at 1 October taken from the true run. This would be the situation in which we have perfect knowledge about the initial state (storage of water in soil, groundwater, lakes, and rivers) of the hydrological system before forecasting.
- Using the average initial conditions. These are obtained by taking the initial hydrological states for 1 October as calculated from PCR-GLOBWB integrations forced by the $P$, $E$, and $T$ of an average hydrological year (i.e., the climatology) as derived from the ERA-40 meteorological data. This amounts to knowing nothing about the initial hydrological state before forecasting.
These two options represent the two end members of an actual forecasting situation. Comparison of two options will provide insight into the added worth of investing in data assimilation schemes to estimate current hydrological states in a seasonal prediction framework.

3) As a reference we compare the skill of NAO-based predictions with predictions using the \(P\), \(E\), and \(T\) of an average hydrological year (i.e., the climatology) as derived from the ERA-40 meteorological data. This is again repeated twice: the exact hydrological state of each year is used as the initial condition; the average hydrological state is used as the initial condition.

Comparing reference runs with the runs described under step 2 will show the skill added when using NAO-based predictions instead of climatology.

Figure 2 provides the four cases/methods that are analyzed: 1) average initial conditions and average forcing; 2) exact initial conditions and average forcing; 3) exact initial conditions and forcing based on NAO-based predictions; and 4) average initial conditions and forcing based on NAO-based predictions.

The skill scores obtained from such a numerical experiment must be viewed as measures of theoretically attainable skill, given the current state of seasonal weather prediction. This means that the experiment will yield measures of skill that are higher than can be expected in reality: First, because both predictions and observations are from the same period (1957–2002), while in a true predictive mode only years from the past can be used for prediction; second, because our “observed” discharges are created with the same model as used to make the hydrological predictions, while in reality a model only approximates the real system. Still, as we will show hereafter, the estimated skill from the numerical experiment compares well with the predictive skill at the 66 gauging stations throughout Europe. Moreover, our main goal is to explore the variation in skill between (stretches of) rivers and understand this variation in relation to river basin properties (e.g., hydrological response time) and the ability to estimate initial hydrological states. A simulation experiment in which “reality” is assumed known everywhere is well suited for such a purpose.

3. Results

a. Verification of PCR-GLOBWB

Our numerical experiment uses a simulated reality to assess the skill of seasonal prediction methods across Europe. To have any confidence in the obtained results, we have to be sure that our model PCR-GLOBWB when forced with ERA-40 data is sufficiently accurate in simulating the discharge of European rivers. Fig. 3 shows a comparison of yearly runoff for the 66 runoff gauging stations from the Global River Discharge (RivDIS) dataset (Vörösmarty et al. 1998). A comparison of average winter (ONDJFM) and summer (AMJJAS) runoff estimated with PCR-GLOBWB and from observations results in variance \(R^2\) of 0.93 and 0.88, respectively.

b. Seasonal prediction of winter discharge

1) Map of skills

Before inspecting the results of the four cases, it is of interest to see whether there are any structural differences in discharge between them. Figure 4 shows maps of average winter cumulative runoff in Europe according to the four cases. The grayscale is proportional to the magnitude of winter-average cumulative runoff, which effectively results in maps of mean winter discharge of the main European rivers. As can be seen, the differences between the four methods are small, or even negligible. Similar results were found for the summer discharge.

Figure 5 shows a map of the skill in seasonal prediction of winter discharge for each of the four cases. Skill is measured as the correlation coefficient between the time series of seasonal predictions and “true discharge”—the latter is the discharge obtained from forcing PCR-GLOBWB with ERA-40. Also denoted is the average skill. The following can be deduced from the maps: First, case 1 provides a skill close to zero everywhere. This is merely a consistency check, as the skill should be zero because we are comparing the same winter discharge for
each year (based on average initial conditions and the winter climatology as forcing) with the actual winter discharge time series. Second, for cases 2 and 3 high skills are achieved for large parts of Europe, suggesting that for many river stretches seasonal prediction may be feasible—in theory. However, when comparing the skills of case 2 to case 3 and case 3 to case 4, it is obvious that the achieved skill mainly comes from knowing the correct initial conditions of the hydrological system (i.e., groundwater, surface water, and soil water storage of the basin). The skill based on the use of NAO-based seasonal weather prediction only (case 4) is rather limited. At 5% accuracy a skill significantly different from zero should be larger than 0.25 \((N = 44)\). This can only be found in limited areas, such as Scandinavia, the Iberian Peninsula, the Balkans, and around the Black Sea.

2) ANALYSIS

We performed further analysis to understand the observed patterns of skill in Fig. 5. In the introduction, it was hypothesized that the slow response time of hydrological systems may be the reason for the large impact of knowing the initial conditions on seasonal prediction skill. Of course, this would only be the case if the initial conditions on 1 October are considerably different between years. If this is the case and the hydrological system reacts slowly to changes in forcing, then a correct estimate of the initial conditions has a large impact on the added skill in predicting the subsequent discharge. Therefore, we calculated (for each pixel separately) two measures to explain the skill due to initial conditions: 1) The standard deviation of soil moisture deficit (difference between actual soil moisture and saturation) on 1 October, as calculated from the 44-yr run of PCR-GLOBWB forced with ERA-40. This measures how different one year is from the other in terms of water that can be stored before runoff occurs. 2) The characteristic response time of the groundwater system as proxy for total hydrological response time. This measure is calculated by taking for each cell (on the 0.5° \(\times\) 0.5° grid) a discharge-weighted average of the groundwater response time of all upstream cells. The groundwater response time is a parameter that links groundwater storage to groundwater drainage (base flow) and was parameterized by van Beek (2007). In periods without precipitation many streams are mainly fed by

![Fig. 3. Comparison of PCR-GLOBWB average yearly discharge based on ERA-40 forcing and observed discharge for 66 stations along major rivers of Europe.](image)
discharging groundwater and show an exponential decrease of discharge with time. The groundwater response time is the e-folding time describing this exponential decrease.

Figure 6 shows the resulting maps of the standard deviation of soil moisture deficit and groundwater response time. In Fig. 7, these values are plotted against the skill achieved by knowing the initial conditions exactly and using the climatology as forcing (case 2). Because two factors are considered, we provide box-and-whisker plots of skill for increasing value classes of the standard deviation of soil moisture deficit and for four classes of groundwater response times. It can be seen that the average skill increases with increasing standard deviation of soil moisture deficit. This indicates that, if the between-year variability of 1 October initial conditions is large, much can be gained from the correct estimation of these conditions. Also, comparing the four plots for the different groundwater response times shows that the sensitivity to the soil moisture standard deviation increases with increasing response time, which is expected because initial conditions matter for a longer time after 1 October if response times are large. Thus for regions where the hydrological system is slow (large groundwater component, lakes or wetlands) and initial conditions are variable, a correct assessment of these initial conditions leads to a large skill in seasonal forecasts of river discharge, even if an average meteorological forcing is used.

To explain the patterns of skill resulting from NAO-based seasonal prediction only (case 4) and how these relate to the skill by knowing the initial conditions, we constructed Fig. 8. The lower-left panel shows again the skill according to case 4 (i.e., using average initial conditions and NAO-based seasonal weather prediction). The upper-left panel shows the added skill by using NAO-based seasonal predictions when initial conditions are known. This figure is obtained by subtracting the map for case 2 from the one for case 3. The upper-right panel shows a map of the NAO sensitivity of winter discharge for European rivers. It shows so-called z scores that are obtained by subtracting, for each river

**Fig. 4.** Maps of average winter discharge for rivers in Europe according to the four cases (modes of seasonal forecasting).
stretch, the average winter discharge of the three years with the smallest winter NAO index from the average winter discharge of the three years with the largest winter NAO index and dividing this result by the between-year standard deviation of winter discharge. Large positive $z$ scores are found in regions with more-than-average river discharge in a NAO-positive year and large negative $z$ scores in areas with less-than-average discharge. So, regions with either large positive or large negative $z$ scores have rivers that are sensitive to the NAO. Figure 8 shows that the added skill by using NAO-based seasonal weather forecasts instead of the climatology, if the initial conditions are known, is largest in the areas where case 4 shows the highest skill. This means that the effects on skill by initial conditions and on skill by NAO-based forecasting are orthogonal—that is, the combined effect is a weighted sum of the two individual effects. Also, the regions with the largest skill resulting from using NAO-based seasonal predictions coincide with the regions where river discharge is sensitive to the NAO phase (large positive or large negative $z$ scores): Spain, the region around the Black Sea (including the lower Danube), and Scandinavia. It can thus be concluded that the highest skill in seasonal prediction can be expected in regions with slowly reacting hydrological systems, large year-to-year variability of hydrological states, and a strong connection between winter NAO phase and winter discharge, provided that the initial state of the hydrological system is estimated with sufficient accuracy.

c. Seasonal prediction of summer discharge

To investigate how long initial conditions at the beginning of winter and NAO-based forecasts remain skillful in time, we also extrapolated the predictions of the four cases to the subsequent summer period. Figure 9
shows the resulting skill maps. The map for case 2 shows that even with average forcing, knowing the initial conditions still results in a significant skill for certain regions in Europe. We performed the same analysis as for the winter discharge (not shown here for brevity) and found regions of significant skill to coincide with regions where between-year variability of initial conditions and groundwater response times are large. A surprising result is that NAO-based prediction without knowing 1 October initial conditions (case 4) results in a slightly higher average skill for summer discharge than for winter discharge. Comparing Figs. 5 and 9 for case 4 shows that higher skill in summer is found for regions where winter precipitation falls partly as snow (e.g., Volga River basin north of the Caspian Sea, central Europe, and the Ukraine). Here, precipitation anomalies in winter result in anomalies in summer runoff. Also, the Iberian Peninsula shows relatively large skill in summer. This can be partly explained because part of the rainfall during the wet period falls in April (part of the summer period in our analysis) but also due to delayed discharge through shallow and complex aquifers (Struckmeier et al. 2004) and large reservoirs (Kilsby et al. 2007). Trigo et al. (2004) also report that the effect of anomalous runoff, related to the NAO, of the three main Iberian rivers (Douro, Tejo, Guadiana) lasts from November to July.

d. Verification with observations

The simulation experiment shed new light on the spatial variation of predictive skill in seasonal discharge forecasts across Europe and how this is related to hydrological properties and NAO sensitivity of the river basins. However, these results are still based on a simulation experiment. By comparing with data, we try to determine whether the predicted skill patterns are realistic.

1) Example time series

Figure 10a shows a comparison of observed winter discharge and the seasonal prediction for four European rivers according to case/method 3. Figure 10b shows the same graphs but for summer discharge. For completeness, we have also included the discharge modeled by the PCR-GLOBWB forward run forced with ERA-40 data (i.e., the true discharge of the simulation experiment). This way one can also see how well the model is able to approximate observed discharge as well as the degree of similarity of predicted discharge and true discharge in the previous simulation experiment. We have selected four rivers as an example. The Angermanälven is the largest and longest river in Sweden and runs from the Scandinavian mountain range to the Gulf of Bothnia (length 470 km, catchment size 31 000 km², average discharge 507 m³ s⁻¹). Its discharge has a mixed regime with a larger snowmelt peak in May and a smaller precipitation peak during winter. The Duero River (Douro in Portugal) is one of the major rivers of the Iberian Peninsula. It runs from central northern Spain westward to the Atlantic Ocean at Porto (length 897 km, catchment size 79 096 km², average discharge at the

![Soil moisture deficit SD (m)](image1)

![Ground water memory (k)](image2)

Fig. 6. (left) Standard deviation of soil moisture deficit (fraction of saturation) and (right) groundwater system response time (months).
Spanish–Portuguese border 563 m$^3$ s$^{-1}$). Its discharge is mainly dependent on winter rains and may be very variable between years. Discharge peaks occur in February and March. The Ångermanälven and the Duero were chosen because they run through regions where winter precipitation is heavily influenced by the NAO (Fig. 8). The Volga and Danube rivers are chosen because they are the largest rivers in Europe and pass through various climate zones. The Danube has its primary source in the German Black Forest. It runs through Austria, Hungary, Serbia, along the border between Romania and Bulgaria, and enters the Black Sea in Romania (length 2860 km, catchment size 817 000 km$^2$, average discharge 6537 m$^3$ s$^{-1}$). Although its main discharge peak occurs in May, various stretches of the river have different runoff regimes: From Figure 8 it can be seen that sensitivity to the NAO is different at different parts of the river. In our analysis, we include the part of Europe (west of the Urals) that is influenced by the NAO. From this part, the Volga is the largest river. Its main tributary has its source in northwestern Russia close to St. Petersburg and runs eastward to Kazan and then southward through Volgograd, after which it enters the Caspian Sea (length 3692 km, catchment size 1 380 000 km$^2$, average discharge at Volgograd 7823 m$^3$ s$^{-1}$). The hydrograph of the Volga has a strong signature of early summer snowmelt with the peak discharge at Volgograd occurring in May. The four rivers considered here are all heavily regulated by dams and reservoirs.

![Box-and-whisker plots of skill achieved by seasonal forecasting of winter discharge for case 2 (initial conditions known; average forcing with winter climatology) for increasing value classes of the standard deviation of soil moisture deficit. A separate plot is given for each class of groundwater response time (four classes as denoted in Fig. 6). Small white squares denote the mean skill for each class, boxes: the interquartile range and whiskers: the minimum and maximum values.](image-url)
If we first compare the forward run of PCR-GLOBWB (straight lines) with observations, it can be seen that, for the selected rivers, the model is able to predict the multiyear variation in winter discharge reasonably well: the Ångermanälven, Wolga, Danube, and Duero with $r = 0.83, 0.45, 0.76,$ and $0.77$ for winter discharge, respectively, and $r = 0.77, 0.70, 0.66,$ and $0.81$ for summer discharge.

For the comparison of seasonal prediction of winter discharge using case 3 (dashed lines) with observations, the skill is plotted in each of the figures. Results for the winter discharge are reasonable in all rivers. Particularly good results are achieved for the Ångermanälven ($r = 0.70$). Here the NAO signal in terms of temperature and precipitation anomalies is strong (Uvo and Berndtsson 2002), while knowing initial conditions also results in large skill scores (Fig. 5). Figure 10b shows that the best results for predicting summer discharge are achieved for the Duero ($r = 0.43$) and the lower Danube ($r = 0.38$). This is because both rivers run through NAO-sensitive regions (the Iberian Peninsula and the area around the Black Sea) and have summer runoff that relies on winter precipitation through groundwater outflow and reservoirs (Struckmeier et al. 2004; Trigo et al. 2004).

2) COMPARISON WITH SKILL BASED ON OBSERVATIONS

To verify the patterns of skill, as displayed in Figs. 5 and 9, we performed the following analysis. We extracted monthly discharge data from the RivDIS dataset (Vörösmarty et al. 1998) for 66 discharge gauging stations along major European rivers with discharge data overlapping the ERA-40 period (1957–2002). For each station we calculated from the monthly data a time series of average winter (ONDJFM) and summer (AMJJAS) discharge. Next, the time series of observed winter and summer discharge was predicted with the three methods of seasonal prediction (cases 2, 3, and 4) and skill scores calculated based on the correlation coefficient between predicted and observed winter/summer discharge. The skill scores thus obtained are the
actual skills based on observations. From our numerical experiments (sections 3b and 3c), we also have the skill scores at the pixels that coincide with the gauging stations. These skill scores are obtained from comparing the seasonal predictions by the three methods (cases 2, 3, and 4) with the simulated reality. If the patterns of skill obtained from our numerical experiment, as shown in Figs. 5 and 9, are correct, then plotting the observed skill at the gauging stations against the simulated skill at the coinciding pixels in Figs. 5 and 9 should yield scatterplots with points that plot around the one-to-one line.

The results of this are shown in Fig. 11. Next to the one-to-one line, we also plotted the regression line and the fraction explained variance $R^2$. Average skills corresponding to cases 2, 3, and 4 for winter discharge are 0.34, 0.31 and 0.08 for winter and 0.09, 0.17 and 0.15 for summer. As expected, the skill in our simulation experiment, on average, overestimates the actual skill based on observations, mainly because the hydrological model error is not taken into account in the experiment. However, we definitely see the same tendency in observed and simulated skill, which indicates that the patterns of skill (as shown in Figs. 5 and 9) are confirmed by the observations. Also, we see the same relationships in skill magnitude between the cases as in the simulation experiment: higher skill on average in winter than in summer, initial conditions more important than NAO-based seasonal prediction for winter discharge, and less so for summer discharge (case 2 versus case 4). These results thus emphasize the conclusions drawn from the simulation experiment.

4. Conclusions and outlook

We performed a simulation experiment using a large-scale hydrological model and the ERA-40 dataset for meteorological forcing to investigate the skill of seasonal prediction of winter and summer discharge for river
stretches across the European continent. Seasonal predictions were based on first predicting the winter NAO index based on May SST (Rodwell and Folland 2002) and then choosing ERA-40 hydrological years (October–September) with similar reported winter NAO index to force a hydrological model for the subsequent year.

We calculated European-wide patterns of theoretically attainable skill (measured in terms of the correlation coefficient between predicted and observed values) of predicting the average winter discharge and subsequent summer discharge. These patterns, corroborated by observations, show that in many parts of Europe the

![Graphs comparing observed winter discharge with predicted discharge for Angermanalven, Volga, Danube, and Duero rivers.](image)

**Fig. 10a.** Comparison with observed winter discharge of the run with PCR-GLOBWB forced by ERA-40 and the seasonal prediction based on case 3: discharge standardized by mean and standard deviation of observed yearly discharge.
skill of predicting winter discharge can, in theory, be quite large. However, this achieved skill mainly comes from knowing the correct initial conditions of the hydrological system (i.e., groundwater, surface water, and soil water storage of the basin) rather than from the use of NAO-based seasonal weather prediction. When predicting further ahead (i.e., subsequent summer discharge), the predictive value of knowing initial conditions decreases, as expected. However, the predictive skill of NAO-based seasonal weather prediction holds its value in many regions of Europe, presumably because these are regions where summer runoff relies on snowmelt, groundwater discharge, or reservoirs.

Further analysis also showed that regions of high skill as a result of knowing initial conditions coincide with regions with a large between-year variation of these
Fig. 11. Predicted skill based on the simulation experiment at the location of 66 RivDIS discharge gauging stations vs the actual skill based on the observations: summer and winter discharge for cases 2, 3, and 4.
initial conditions and a large system memory, which indicates that a correct estimate of the initial state has a large predictive value for the discharge in the months to come. As expected, regions where NAO-based seasonal weather predictions lead to reasonable skill are those regions where winter precipitation is highly correlated with NAO phase, such as Scandinavia, the Iberian Peninsula, and the regions around the Black Sea.

From our study one may arrive at the conclusion that, given the limited skill at which we are currently able to predict the NAO index of the coming season, seasonal forecasting of river discharge requires the development of accurate nowcasting systems using good models, new observational techniques (e.g., remote sensing), and data assimilation to estimate current hydrological states (i.e., storage in groundwater, rivers, lakes, wetlands, and, for summer discharge, snowpack). We stress that this conclusion probably holds only for Europe at the present time. First, as the understanding of European large-scale circulation dynamics increases, seasonal prediction can be expected to become better as well, for instance, by using multimodel ensembles of very high-resolution OAGCMs (Rodell and Doblas-Reyes 2006). Second, seasonal weather prediction is known to be more successful in other parts of the world. For instance, the strength of the monsoon at lower latitudes may well be better predictable than the NAO across Europe, as it is generally more strongly related to SST than the NAO (e.g., Van den Dool et al. 2006). Also, multiyear circulation phenomena that are more ocean driven, such as ENSO, are better predictable and will probably lead to large predictive skill of river discharge in many areas of the world. A similar analysis weighing hydrological nowcasting against seasonal weather prediction would be a worthwhile undertaking for these areas.

Finally, the study presented here is only a first attempt to consider seasonal predictability of discharge in relation to hydrological state estimation and seasonal weather prediction. A number of aspects deserve more attention. First, instead of only investigating the cases of no knowledge and perfect knowledge about initial hydrological states, it is worthwhile to see how the accuracy at which initial hydrological states can be estimated relates to the skill of seasonal forecasts. Here, the use of various data-assimilation methods and remote sensing products to estimate hydrological states becomes an issue (Rodell et al. 2004). Second, we have only considered a coarse temporal resolution, that is, season averages. A more in-depth analysis of predictability would be obtained by looking at predictability at monthly time scales for combinations of lead times and time of year at which the prediction is made. Finally, related to that is the question how the frequency and timing of initial hydrological state estimation impacts predictability. For instance, in areas with a large contribution of snowmelt to runoff a correct estimate of snow cover and snow depth at the end of the cold season would most likely be the prime factor explaining predictive skill (Wood and Lettenmaier 2006).

Acknowledgments. We would like to acknowledge the help of three anonymous reviewers whose comments improved this paper considerably.

REFERENCES


Struckmeier, W. F., W. H. Gilbrich, A. Richts, and M. Zaepeke, 2004: WHYMAP and the groundwater resources of the world at the scale of 1:50 000 000. [Available online at http://www. whymap.org.]


——, C. S. Wedgebrow, and H. R. Fox, 2004: Seasonal predictability of the summer hydrometeorology of the river Thames, UK. *J. Hydrol.*, **295**, 1–16.
