STATISTICAL METHODS FOR RIVER’S RUNOFF FORECAST

Pisarenko V.F.¹, Lyubushin A.A.², Canu S.³, Kanevsky M.F.⁴, Savelieva E.A.⁴,
Demianov V.V.⁴, Bolgov M.V.⁵, Rukavishnikova T.A.¹, Zaliapin I.V.¹

¹International Institute of Earthquake Prediction Theory and Mathematical Geophysics, RAS, Moscow; ²Institute of
Physics of the Earth, RAS, Moscow; ³Institut National des Sciences Appliquees, Rouen, France; ⁴Nuclear Safety
Institute, RAS, Moscow; ⁵Water Problem Institute, RAS, Moscow.

ABSTRACT

Statistical methods of data analysis are applied to monthly runoff values of rivers. The analysis
is based on the theory of periodically correlated random processes. This approach takes into account
both deterministic and stochastic components of seasonal variations of runoff data. Prediction
formulae of runoff value one month ahead, based on 1-12 preceding monthly runoff values, are
derived. This method was applied to 9 rivers with different hydrological regimes. The efficiency of
prediction, which can be defined as ratio of residual variance to initial variance of runoff, varies
from river to river in the range $0.1 \div 0.6$. The suggested method of prediction was compared with
Artificial Neural Networks. Factor analysis of correlation matrices of deviations of rivers runoff
from the cyclic average allowed extracting the maximum number of common orthogonal factors.
This number gives the measure of complexity of the runoff dynamics. Hydrological aspects of
results of statistical analysis of runoff are discussed.

Key words: runoff forecasting, periodically correlated random processes, artificial neural
networks.

INTRODUCTION

River’s runoff forecasting is a difficult hydrological topic, since underlying physical processes
are complex and far from being considered as adequately described by a system of the
corresponding equations. Therefore, statistical and “data mining” methods are widely used aiming
to forecast rivers’ runoffs (Fernandez, Salas, 1974; Moss, Bryson, 1974; Vecchia, 1985; Loucks,
Stedinger, Haith, 1988; Privalsky, Panchenko, Assarina, 1992). Several modifications of statistical
forecast based on the autoregression time-series analysis combined with harmonic trend modeling
were suggested in (Kashyap, Rao, 1976). But they all have an essential deficiency: the stochastic
component is modeled by stationary autoregression whereas it exhibits a distinct seasonal
periodicity.

The reliable runoff forecast is necessary for a large number of applications connected both with
water use (water-supply, power, navigation, etc.) and with protection from flooding. Certainly, the
reliability of operative runoff forecast depends in many respects on existing monitoring systems
providing necessary data for prediction, and such systems should be constantly developed. The
problem described in this article consists in developing forecast methods based only on past runoff
data. Hydrological experience shows that in such formulation long-term forecasts are hardly
possible. In contrast, forecast of one time step ahead (one month ahead) is quite realistic which
permits more effective use of regulating capacities of water reservoirs. Theoretical results of the
paper include conclusions on type and complexity of stochastic runoff model with sampling interval
one month (in other words - model with seasonal course). Such models are of particular interest in
investigations of complex water economic systems using simulation methods. Examples of models
with seasonal course are known in hydrology (see e.g. (Ratkovich, Bolgov, 1997) and references
therein), but their properties and quality of forecast furnished by them have been studied quite insufficient.

In this study, one comparatively new statistical technique, namely, the theory of periodically correlated processes (PCP) (Rytov, 1976) is used for this purpose. Sometimes this model is referred to as cyclic stationary process (Gardner, Franks, 1975). Prediction formulae of runoff value one month ahead, based on the preceding monthly runoff values are derived with the help of standard least square technique. The PCP approach takes into account both deterministic and stochastic components of seasonal variations of runoff data. The method of PCP permits to adjust prediction coefficients to particular month of issuing a forecast. The prediction coefficients vary from one month to another. We used artificial neural network (ANN) runoff predictions as a benchmark of forecasting efficiency (Cheng, Titterington, 1994; Haykin, 1999). This approach was applied for rivers runoff forecasting for instance in (Jayawardena, Fernando, 1996; Dawson & Wilby, 1998; Campolo, Soldati, Andreussi, 1999; Jayawardena et al., 2000). A perceptron neural network with one hidden layer trained on the past time period was used for a monthly forecast of the runoffs. The results of comparison of the two approaches show that the PCP prediction is never worse than ANN predictions. As a rule, the PCP is by 10-30% more efficient than the ANN, but for some rivers the PCP and ANN are practically the same.

We have used runoff data of 9 large rivers of Europe and Asia: Loire, Elba, Danube, Glomma, Wisla, Oka, Northern Dvina, Irtysh, Amur. Time intervals of runoff registration were different for these rivers ranging from 84 up to 134 years.

THE DATA

We have chosen 9 large rivers of Europe and Asia for demonstrating our prediction method: Loire (1863-1979), Elba (1851-1984), Danube (1901-1984), Glomma (Norway, 1901-1984), Wisla (1900-1987), Oka (Central Russia, 1881-1985), Northern Dvina (Northern Russia, 1881-1985), Irtysh (West Siberia, 1891-1985), Amur (Far East of Russia, 1896-1985). The data consisted of monthly runoff values \( X(t), t=1,\ldots,N \). The general view of runoff time series is shown on Fig. 1. These time series differ from one another rather significantly. Some rivers (Glomma, Amur, Irtysh, Oka, N.Dvina) show clear regularity connected with the annual maxima caused by snow melting and spring floods. These time series are typical examples of PCP. Other rivers (Wisla, Elba, Loire) are less regular. Danube features the most chaotic behavior. It can be explained by smaller contribution of spring floods in water regime of these rivers.

Average monthly values of runoffs (year’s forms) \( m(\beta), \beta=1,\ldots,12 \), (further on letter \( \beta \) will be used only as the index for enumerating months) are shown in Fig. 2 along with the corresponding standard deviations. They differ from one another as well. Some of them (Loire, Danube, Elba) demonstrate a regular, quasi-harmonic behavior, that can be explained by weak spring floods and human impact in the corresponding water regime. Other rivers (Glomma, Wisla, Oka, N.Dvina, Irtysh) show one distinct peak, corresponding to spring floods. Annual dynamics of Amur has got two peaks: the smaller one corresponds to spring flood, whereas the larger one – to autumn floods caused by cyclones. Regards standard deviations in Fig. 2, we can emphasize one feature common to all rivers: they are proportional to monthly runoff values, which is quite understandable. It should be noted, that relative standard deviations are larger for months with high runoff values as well.

Power spectra of runoff time series are shown in Fig. 3. A distinctive feature of all these spectra is sharp peaks that are separated by equal intervals. These peaks are caused by periodic deterministic components with a period of one year and its overtones. The peak amplitudes are closely connected with the corresponding annual shapes, since power spectra are smoothed Fourier coefficients of the annual shapes. Thus, \( \delta \)-like annual shapes for Oka and N.Dvina give rise to 6
spectral peaks of almost equal amplitudes, whereas harmonic-like year forms of Loire, Danube, Elba give one prominent spectral peak.

**PERIODICALLY CORRELATED RANDOM PROCESSES (PCP)**

We recall main definitions of PCP (for more details see (Rytov, 1976; Gardner, Franks, 1975)). A random process \(X(t)\) with discrete time \(t\) is called periodically correlated random process with period \(T\), if its mean value: \(m(t) = E\{X(t)\}\) and covariance function: \(C_{xx}(t,s) = E\{(X(t)-m(t))(X(s)-m(s))\}\) are periodic functions with period \(T\): \(m(t+T) \equiv m(t)\); \(C_{xx}(t+T,s+T) \equiv C_{xx}(t,s)\).

In our case of runoff data \(T=12\). As it was mentioned above, runoff time series demonstrate an evident annual seasonal periodicity. Of course, this fact can be used for the prediction. Thus, the annual shapes in Fig. 2, which are in fact mean functions \(m(t)\), represent simple, deterministic forecast of monthly runoffs. But more interesting and valuable problem is to forecast the stochastic component of the runoff time series. It should be noted that in order to satisfy “the causality principle” it is necessary to construct an estimate of \(m(t)\), which is used in a prediction issued at the moment \(t\), based only on the past values \(X(s), s < t\).

Centered time series \(Y(t)\) are shown on Fig.4. It is natural that after subtraction of means sample variances \(\delta_Y^2\) of \(Y(t)\) become less than sample variances \(\delta_X^2\) of the original time series \(X(t)\). Ratio \(\gamma_m = \delta_X^2/\delta_Y^2\) can be considered as an index of deterministic predictability, when prediction is based only on the cyclic mean \(m(t)\). The values \(\gamma_m\) are shown in Table 1. We see, that subtraction of cyclic mean \(m(t)\) reduces sample variance rather significantly: from 1.09 up to 5.21 times. However, this decreasing is not uniform for different months. The PCP-model just permits to use this heterogeneity.

Thus, the most deterministically predictable rivers have appeared to be those with strong seasonal floods: Irtysh, Amur, N.Dvina, Oka. The worst predictable are the rivers with weak spring floods: Danube, Elba, Loire, Wisla. But this ordering will be changed when we consider the predictability of the stochastic component \(Y(t)\) (see below).

Tables 2a and 2b give examples of correlation matrices for two rivers – Oka and Irtysh. It is seen that Oka has rather low correlation coefficients for neighboring months, whereas Irtysh has larger coefficients going up to 0.92 (February-March). The PCP-model just takes into account changes of correlation coefficients from one month to another. Now we are going to derive prediction formulae for PCP-model.

**PREDICTION OF RUNOFF**

Prediction value \(\hat{X}(t)\) of the process is sought in the form: \(\hat{X}(t) = \hat{m}(t) + \hat{Y}(t)\). Here \(\hat{m}(t)\) is an estimate of the cyclic mean \(m(t)\) based on the past values \(X(s), 1 \leq s < t\). Prediction \(\hat{Y}(t)\) of the stochastic component \(Y(t)\) is sought in the form of a linear combination of \(p\) past observations:

\[
\hat{Y}(t) = a_1(t)Y(t-1) + \ldots + a_p(t)Y(t-p)
\]  

where \([t-p+1] \ldots t\) is the time interval used for prediction. It should be noted that prediction coefficients \(a_j(t)\) do depend on the time moment \(t\) of the issuing prediction. The question of choice of the order \(p\) in the model (1) is discussed below.
We have tested two variants of estimating the coefficients $a_j(t)$ in (1). The first one we call \textit{standard AR-estimating} (Box, Jenkins, 1970; Kashyap, Rao, 1976) and it consists of the following procedure. Let $L$ be a number of readings within used time window for the current time moment $t$ of issuing prediction $\hat{Y}(t)$. Then the standard AR-estimating corresponds to coefficients of linear prediction minimizing the sum of squared errors of prediction:

$$\sum_{n=1}^{L-p} (Y(t-n) - \sum_{k=1}^{p} a_k \cdot Y(t-n-k))^2 \rightarrow \min_{a_k}$$

(2)

It should be noted that unlike the PCP approach the prediction coefficients $a_k$ in eq.(2) are not assumed to be dependent on the time moment $t$.

Solution of problem (2) should satisfy the system of linear equations:

$$\sum_{k=1}^{p} C_{k,j}^{(0)}(t) \cdot a_k = C_{0,j}^{(0)}(t), \ j=1,\ldots,p$$

(3)

where:

$$C_{k,j}^{(0)}(t) = \sum_{n=1}^{L-p} Y(t-n-k) \cdot Y(t-n-j)$$

(4)

Let us denote the solution of the linear system (3) as $a_k^{(0)}(t \mid p)$. We have left in this notation a dependence of $a_k$ on $t$. It should be noted, that for infinite ensemble averaging $(L \rightarrow \infty)$ this dependence vanishes for stationary time series $Y(t)$. But for a finite $L$ some dependencies of $a_k$ on $t$ can remain.

The second variant, that we call \textit{PCP AR-estimating}, corresponds to the modification of the minimum least squares problem (2), which takes into account property of the periodic correlation. Let $q$ be an integer part of the ratio $(L-p)/T$: $q=[(L-p)/T]$. The integer number $q=q(t)$, generally speaking, depends on $t$. Now, we have to solve the problem:

$$\sum_{n=1}^{q(t)} (Y(t-n \cdot T) - \sum_{k=1}^{L-p} a_k \cdot Y(t-n \cdot T-k))^2 \rightarrow \min_{a_k}$$

(5)

From (5) we derive the following linear system:

$$\sum_{k=1}^{p} C_{k,j}^{(1)}(t) \cdot a_k = C_{0,j}^{(1)}(t), \ j=1,\ldots,p$$

(6)

where:

$$C_{k,j}^{(1)}(t) = \sum_{n=1}^{q(t)} Y(t-n \cdot T-k) \cdot Y(t-n \cdot T-j)$$

(7)

We denote the solution of system (6) by $a_k^{(1)}(t \mid p)$. Thus, we have two different linear predictors of the order $p$:

$$\hat{Y}^{(\alpha)}(t \mid p) = \sum_{k=1}^{p} a_k^{(\alpha)}(t \mid p) \cdot Y(t-k), \ \alpha=0,1$$

(8)
If we compare method (2) with (5) it can be easily seen that the sum (2) contains \( T=12 \) times more members than (5), providing, seemingly, more reliable averaging. However, if the time series under investigation is not stationary, but, on the contrary, is essentially periodically stationary, then such averaging can deteriorate prediction, whereas the averaging (7), taking into account the main property of periodicity of PCP, can lead to more efficient prediction. So, the advantage of PCP prediction depends on what extent time series under question is really periodically correlated, i.e. its correlations vary periodically. We are going to try both these AR-methods and to combine them in order to reduce the prediction error.

For both methods we can construct time series of the prediction errors:

\[
e^{(\alpha)}(t \mid p) = Y(t) - \hat{Y}^{(\alpha)}(t \mid p), \quad \alpha=0,1; \quad t=L_0+1,\ldots,N
\]  

(8)

Here \( L_0+1 \) is the first time moment for prediction, \( N \) is the number of readings in the considered time series. The first \( L_0 \) readings are used for initialization of the algorithm. For PCP AR-estimating it would be convenient to start numeration of all time series with January. Since some time series do not satisfy this condition (see Fig. 1), we have cut off the first “superfluous” readings (not more than 11 first readings), so that the first reading corresponds strictly to January for all series. Besides, we take the length of initialization \( L_0 \) or the length of moving time window \( L \) to be equal to an integer number of years: \( L_0 \) or \( L=L-Tn_0=12-n_0 \), where \( n_0 \) is an integer number. Further, in all cases we use \( n_0=30 \) years, \( L=L_0=360 \) readings. Thus, the time series of prediction errors (8) also starts by January. We denote by \( N_0 \) the integer part of \( N/12 \), i.e. the number of full years of observation. Note that time series \( Y(t) \) can be defined only starting from the 2\textsuperscript{nd} year of observation.

The time series \( Y(t) \) can be split into 12 “sub-series”:

\[
Y(\tau \mid \beta) = Y(12 \cdot (\tau - 1) + \beta), \quad \beta=1,\ldots,12; \quad \tau=2,\ldots,N_0
\]  

(9)

where index \( \beta \) enumerates months and index \( \tau \) enumerates successive years of observation. Thus, \( Y(\tau \mid \beta), \tau=2,\ldots,N_0 \), is the time series of deviations from the cyclic mean value, corresponding to the month with number \( \beta \). Analogously to the formula (9) we can split the time series of prediction errors (8):

\[
e^{(\alpha)}(\beta \mid \tau) = e^{(\alpha)}(12 \cdot (\tau - (n_0 + 1)) + \beta \mid p), \quad \tau=n_0+1,\ldots,N_0
\]  

(10)

We can introduce a value that characterizes the efficiency of the prediction for the month \( \beta \):

\[
\mu^{(\alpha)}(\beta \mid p) = \sum_{\tau=n_0+1}^{N_0} Y^2(\tau \mid \beta) / \sum_{\tau=n_0+1}^{N_0} (e^{(\alpha)}(\tau \mid p))^2
\]  

(11)

Besides, we can introduce two values characterizing total efficiency of the prediction, averaged over all months:

\[
\gamma_0 = \sum_{t=L_0+1}^{N} X^2(t) / \sum_{t=L_0+1}^{N} e^2(t), \quad \gamma_1 = \sum_{t=L_0+1}^{N} Y^2(t) / \sum_{t=L_0+1}^{N} e^2(t)
\]  

(12)

For linear AR-predictors we tried both moving and increasing time windows. Increasing time windows always give a better result. That is why we present below the results for increasing time windows only.
Table 3 contains results of our experiments with AR-prediction for all 9 time series. For each series we tested prediction algorithm produced by AR with $p$ varying from 1 up to 12. We used increasing time window, starting from the 361-st reading (the January following the first 30 years used for initialization). Results of the prediction with the standard AR-estimation ($\alpha=0$) and by PCP AR-estimating ($\alpha=1$) were compared. For each month $\beta=1,\ldots,12$ and for each method $\alpha=0,1$ we calculated month’s prediction efficiency $\mu^{(\alpha)}(\beta|p)$. Let $\alpha^*(\beta)$, $p^*(\beta)$ be solution of simple maximization problem:

$$\alpha^*(\beta), p^*(\beta) : \mu^{(\alpha)}(\beta|p) \rightarrow \max_{\alpha \cdot p}$$  \hspace{1cm} (13)

We call sequence of pairs $(\alpha^*(\beta), p^*(\beta))$ prediction scenario because it provides a rule of choice AR-estimation type $\alpha$ for each AR-order $p$ and month $\beta$. Table 3 contains prediction scenario for each river. The lines marked with the letter $\mu$ contain prediction efficiencies for each month, and total efficiencies $\gamma_0$ and $\gamma_1$ as a result of using particular prediction scenario. The last column in Table 3 contains values of $\gamma_m$ from the Table 1 – for comparing it with $\gamma_0/\gamma_1$. It should be noticed that $\gamma_0$ equals to the product $\gamma_m\gamma_1$ only approximately. This is a consequence of the fact, that ratio $\gamma_m$ was calculated for time interval $[\tau+1,N]$, whereas (12) – for time interval $[L_0+1,N]$. Value $\gamma_1$ gives prediction efficiency for the stochastic component.

It should be noted that scenarios (13) in Table 3 were obtained by processing the whole intervals of observations for each river. So these scenarios could be useful for prediction in the future and we have estimates of their prediction efficiencies. At the same time it would be interesting to check prediction algorithms by standard validation technique: to divide the interval of observation into two equal parts, to use the first half for seeking scenario and the second half for its examining. The main obstacle for such experiments is small duration of time intervals of observations. The duration (i.e. the number of used samples) for PCP AR-estimation is the number of years $N_Y$. For initialization (i.e. for computing AR-coefficients) of AR-predictor we use the initial time interval of the length $n_0=30$ years, which is scarcely sufficient for reliable PCP AR-estimation, in particular for big values of $p$. That is why we could not diminish $n_0$. But if we divide the interval into two parts, then we’ll have only $(N_Y/2-n_0)$ samples for seeking scenario at the first half and for examining its quality at the second half. Taking into account available values of the duration $N_Y=84-134$ years, we have got 12-37 readings for the last number instead of 54-104 $(N_Y-n_0)$ for the results shown in Table 3.

Nevertheless, we have performed validation experiment for 5 rivers with longest intervals of observations: Irtysh, N.Dvina, Oka, Elba, Loire. The intervals of observations were divided into 2 approximately equal parts – not strictly equal because the first and the second parts should start from January. Table 4 contains results of validation.

First of all, application of prediction scenarios determined by the first half of observation interval to the second half gave similar results for all months and for all total efficiencies. Comparing Tables 3 and 4 we notice that scenarios and efficiencies can sometimes differ significantly (as the consequence of small number of samples) but the efficiencies are almost the same, in particular the total efficiencies. This fact demonstrates certain stability of the prediction results shown in Table 3.
ARTIFICIAL NEURAL NETWORK MODEL

This section is devoted to application of Artificial Neural Network (ANN) technique to the problem of prediction of time series $Y(t)$ – stochastic component of rivers’ runoff. We are going to carry out the ANN-prediction in a moving time window of length $L$ and compare it with usual AR-prediction in the same time window. We use the length $L$ of the time window, which is equal to 288 readings (24 years). The necessity of using moving time window follows from computational complexity in calculation of ANN-parameters for increasing time window (for moving window we need only “slight improvement” of current ANN-parameters). For comparison we use only standard AR-estimating with constant AR-order $p=12$. We did not use PCP AR-estimating for that case because of small number of readings $(288/12=24)$, which is not sufficient for reliable statistical estimation of AR-coefficients. At the same time for standard AR-estimating 288 readings is quite enough. Besides, we omit time-consuming procedure of choosing the best AR-orders and take the maximum AR-order $p=12$. Such choice does not make AR-prediction substantially worse.

We have used a simple ANN architecture with one hidden layer, consisting of $p$ neurons. Let us try to make prediction for time moment $t$. For this purpose we consider $q$ preceding values of the time series, compute their weighted sum and use this scalar value as an input to each neuron. Thus, this procedure is based on forming $p$ scalar signals:

$$z_j(t) = \sum_{i=1}^{q} w_{ji} \cdot Y(t-i) + c_j, \quad j=1,\ldots,p$$

(15)

Each signal (15) is passed through the non-linear activation function $f(s)$. We used the same function for each neuron:

$$\xi_j(t) = f(z_j(t)), \quad f(s) = s/(|s|+1/2)$$

(16)

$\xi_j(t)$ are output neuron signals. Prediction for the value $Y(t)$ is computed as a weighted sum of output neuron signals plus a shift parameter:

$$\hat{Y}^{(n)}(t) = \sum_{j=1}^{p} \alpha_j \cdot \xi_j(t) + \beta$$

(17)

Upper symbol “$n$” in the formula (17) indicates that this is ANN-predictor. Thus, the full vector of ANN parameters is the following:

$$\theta = (\beta, \alpha_j, c_j, w_{ji}, \quad j=1,\ldots,p; \quad i=1,\ldots,q)$$

(18)

Vector $\theta$ has dimensionality $pq + 2p + 1$. So predictor (17) can be written in the following form: $\hat{Y}^{(n)}(t) = \hat{Y}^{(n)}(t | \theta)$. Vector of parameters $\theta$ is determined from minimizing a quadratic cost function:

$$J(\theta) = \sum_{t=L+q+q}^{+} (Y(t) - \hat{Y}^{(n)}(t | \theta))^2 \rightarrow \min \theta$$

(19)
where $\tau$ is set to $L$ for the first time window of initialization. Further on $\tau$ will be the time moment corresponding to right-hand end of the current moving time window, $\tau=L,\ldots,N$. We denote by $\tilde{\theta}(\tau)$ vector of parameter estimates obtained by solving the problem (19).

It should be noted that function $J(\theta)$ has, as a rule, large number of local minima due to high dimensionality of vector $\theta$ and non-linear character of this function. Attempts to find the global minimum are not always practicable. In order to find an appropriate minimal value of (19) we took $10^3$ random initial values of $\theta$ out of some parallelepiped containing zero point with half-length of its side $10^{-3}$. For each initial value $\theta$ we started a gradient method with maximum gradient step $10^{-4}$ of searching local minimum with decreasing by $\frac{1}{2}$ the gradient step in cases of increasing the cost function. The total number of steps of the gradient method was limited either by $10^4$ steps, or by the rule that the step value becomes less than $10^{-8}$. Finally the total minimum was selected out of all achieved local minima.

After obtaining the estimate $\tilde{\theta}(\tau)$ for the first time window, this window starts to move from left to right with the step 1 and the initial estimate of vector of parameters is currently innovated by the gradient procedure, taking previous vector $\theta$ as initial point. Thus, for each time window we can define one-step-ahead prediction: $\hat{Y}_F^{(n)}(\tau+1) = \hat{Y}_F^{(n)}(\tau+1 | \tilde{\theta}(\tau))$ for $\tau=L,\ldots,N-1$, the error of ANN-prediction:

$$\delta(\tau) = Y(\tau) - \hat{Y}_F^{(n)}(\tau), \tau = (L+1),\ldots,N$$

and the sample estimate of its variance in the moving time window of the same length $L$ depending on right-hand end of this window:

$$\tilde{\sigma}_F^2(\tau) = \sum_{t=\tau-L+1+q}^{\tau} (\delta^{(n)}(t))^2 / (L-q), \tau=2L,\ldots,N$$

For ANN-predictor we have used $L=288$, $p=2$, $q=3$, the whole number of ANN-parameters equals 11 – approximately the same as for AR-predictor.

Now we can compare 4 estimates of variances in moving time window:
- for initial time series $X(t)$;
- for stochastic component $Y(t)$;
- for errors of prediction by standard AR-estimating with AR-order $p=12$;
- for errors of ANN-prediction (formula (21)).

Fig.5 contains these 4 graphs for 4 rivers: Irtysh, Amur, Danube and Elba. It can be noticed that both AR- and ANN-predictors gain decreasing of variance but AR-predictor is the best one for all cases. The same conclusion is true for other rivers. Thus, our results do not confirm numerous statements of the ANN-enthusiasts about advantage of the ANN-approach to the prediction of time series. The simpler and faster classical methods have appeared to be more effective too. We are ready to take part in discussion on this subject and are ready to give our data to anyone who would wish to obtain better prediction by an ANN approach.

**FACTOR ANALYSIS OF CORRELATION MATRICES**

In this section we present some results concerning factor analysis of correlation matrices presented in Table 2a and 2b, i.e. matrices of the correlation coefficients of $Y(\tau | \beta)$. The purpose
of this analysis is to find some measure of complexity of periodically correlated stochastic signals \( Y(t) \). According to the general approach of factor analysis we write \( Y(\tau | \beta) \) in the following form:

\[
Y(\tau | \beta) = \sum_{k=1}^{m} \lambda_{\beta k} \cdot f_k + \varepsilon(\tau | \beta)
\]

(22)

where \( \lambda_{\beta k} \) is factor loading matrix, \( f_k \) are orthogonal common factors, \( \varepsilon(\tau | \beta) \) are residual values, \( \tau \) is index for numerating years of observations. In order to find factor loading matrix, we use the method of minimum residuals (Harman, 1967).

The main problem in factor analysis is the determination of the optimal value of \( m \). We simplify this problem and test all numbers of \( m \) starting with the minimum \( m=1 \), then increasing it by 1, until the factor analysis model (22) is degenerated. The last means that the variance estimate of some residual becomes much less (by several orders) than variances of other residuals components. In factor analysis this situation is known as Heiwood’s case. It means that number of common factor is too large, and the model (22) is over-fitted. We denote by \( \kappa \) the maximum number \( m \) of common orthogonal factors, which still does not cause over-fitting. We regard this value as the index of complexity of the signal \( Y(t) \) – the more is \( \kappa \), the more complex is \( Y(t) \). Table 5 contains the values of \( \kappa \) and values of total efficiencies from the Table 3 for all rivers.

Figure 6 contains graphs of each of the series \( Y(\tau | \beta) \) for Amur and graphs of two common orthogonal factors (after performing “varimax” factors’ rotation (Harman, 1967)). It could be noticed that the first factor is the most similar to February and the second – to November. This is confirmed by estimating the correlation coefficients between common factors and all other months. So, these two months could be called “key” months for understanding the hidden periodic structure of the hydrological regime of Amur.

**DISCUSSION AND CONCLUSIONS**

The statistical analysis carried out in this work allows determination of some parameters characterizing the peculiarities of particular hydrological regime. First of all, the index of predictability by cyclic mean should be mentioned (see Table 1). All considered 9 rivers can be divided by value of this index into two groups: with high efficiencies \( \gamma_m \) (N.Dvina, Amur, Oka, Irtysh, Glomma) and with low ones (Loire, Wisla, Elba, Danube). The first group is characterized by intensive seasonal floods. Such rivers can be called “seasonally conditioned”. On the contrary, the second group consists of rivers, whose regime depends weekly on snow melting spring floods.

The next characteristic of a river regime is given by the shape of its spectra (see Fig.2). The number of harmonics in the power spectra varies for different rivers. They can be divided by this characteristic into three groups. The first group has one essential harmonic (Loire, Danube). Perhaps, Elba can be referred to the first group too, although a week second overtone is seen on the graph of power spectrum. One prominent spectral peak means that the trend (cyclic mean) looks very close to a harmonic. Indeed, we can verify it looking at Fig.3. At the other extreme, power spectrum can contain maximum possible number of peaks (6 for our case, when 12-months cyclic mean is expanded into Fourier series). Besides, all these spectral peaks are almost of the same magnitude. This group of rivers includes Oka, N.Dvina and, perhaps, Irtysh and Glomma. One peak means, that cyclic mean looks somewhat like \( \delta \)-function, i.e. among 12 monthly runoffs there is one dominating all other. We can verify this fact looking at Fig.3. At last, there are “intermediate” number of peaks in FT of cyclic mean. Wisla and Amur have such features. This characteristic of river regime can be called “spectral complexity of cyclic mean”. Kashyap and Rao modeled the monthly cyclic mean in their examples (Kashyap, Rao, 1976) by a trigonometric polynomial of 2 \( \div \)
4 order. So, these examples correspond to intermediate “complexity”. As we see from our examples other situations with “complexity” number are quite possible. In order to try to characterize “complexity” of runoff regime from another point of view, we applied factor analysis to our data. The number $\kappa$ gives another measure of complexity of runoff dynamics: the more is the number of common factors the more complex is the behavior of the runoff time series. It turns that this number varies from 1 (for Irtysh) up to 5 (for Elba and Danube). Intermediate values of the common factors number are the following: 2 – for Amur, Glomma and Wisla. 3 – for N.Dvina and 4 – for Oka and Loire (see Table 5). For some months the behavior of common factors is similar to the behavior of deviations from the mean value. This we can extract some “key” months in the runoff dynamics.

It should be kept in mind that the latter “complexity” refers to stochastic component of runoffs, whereas the former “complexity” refers mainly to cyclic mean. They are not always correlated (compare, e.g two “complexities” for Oka), but often they exhibit a qualitative correlation (see e.g. Danube, Elba, Loire). Examples of “key” months, which are closely correlated with an unobserved, hidden factor, are shown on Fig.6.

In general the predictability of monthly runoffs achieved by PCP-model and other statistical characteristics studied in this work provides useful information both for theoretical study of hydrological regime of rivers and for some practical usage of this information.

ACKNOWLEDGEMENTS

This work was supported by INTAS grant 99-00099, NSF grant EAR-9804859 and ISTC grant 1293-99.

REFERENCES


FIGURE CAPTIONS.

Fig.1. Monthly water runoffs of 9 rivers of Russia and Europe, N is number of readings in each series.

Fig.2. The year forms (monthly average values) of runoff \( m(\beta), \beta=1,\ldots,12 \) with standard deviation bars.

Fig.3. Power spectra estimates for monthly water runoff.

Fig.4. Deviations from the current estimates of the cyclic mean (signals \( Y(t) \)) for monthly water runoff.

Fig.5. Sample estimates of variance in a moving time window (L=288, 24 years) for 4 rivers. The upper curve is the variance of initial time series \( X(t) \) (see Fig.1). The lower curves are different estimates of variance of stochastic component \( Y(t) \) (see Fig.4): dashed line is estimate of ANN-prediction error (eqn. (21)); the lowest curve is estimate of standard AR-prediction error in moving time window with AR-order \( p=12 \). All curves are given as functions of right-hand end of moving time window. Mark on time axis corresponds to the number of months since the beginning of observations.

Fig.6. Deviations from the cyclic mean for Amur, months I-XII and graphs of 2 common orthogonal factors for deviations.
### Table 1. Indices of deterministic predictability (prediction is based on cyclic mean only).

<table>
<thead>
<tr>
<th>River</th>
<th>N.Dvina</th>
<th>Amur</th>
<th>Oka</th>
<th>Irysh</th>
<th>Glomma</th>
<th>Loire</th>
<th>Wisla</th>
<th>Elba</th>
<th>Danube</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_m$</td>
<td>5.21</td>
<td>4.38</td>
<td>2.95</td>
<td>4.15</td>
<td>3.39</td>
<td>1.47</td>
<td>1.39</td>
<td>1.27</td>
<td>1.09</td>
</tr>
</tbody>
</table>

### Table 2a. Matrix of correlation coefficients between different months for deviations from the cyclic mean ($Y(t)$) for Oka.

<table>
<thead>
<tr>
<th>β</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.57</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.08</td>
<td>0.19</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.07</td>
<td>-0.13</td>
<td>-0.51</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.00</td>
<td>-0.07</td>
<td>-0.31</td>
<td>0.27</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.12</td>
<td>0.06</td>
<td>0.39</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.10</td>
<td>0.06</td>
<td>0.11</td>
<td>0.02</td>
<td>0.31</td>
<td>0.68</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.10</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.11</td>
<td>0.16</td>
<td>0.44</td>
<td>0.63</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.13</td>
<td>0.03</td>
<td>-0.01</td>
<td>-0.07</td>
<td>0.10</td>
<td>0.48</td>
<td>0.48</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.16</td>
<td>0.02</td>
<td>-0.10</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.17</td>
<td>0.23</td>
<td>0.40</td>
<td>0.57</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.12</td>
<td>0.01</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.05</td>
<td>0.22</td>
<td>0.28</td>
<td>0.45</td>
<td>0.59</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.19</td>
<td>0.09</td>
<td>0.03</td>
<td>-0.12</td>
<td>-0.08</td>
<td>0.00</td>
<td>0.11</td>
<td>0.21</td>
<td>0.20</td>
<td>0.37</td>
<td>0.51</td>
<td>1.00</td>
</tr>
</tbody>
</table>

### Table 2b. Matrix of correlation coefficients between different months for deviations from the cyclic mean ($Y(t)$) for Irysh.

<table>
<thead>
<tr>
<th>β</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.70</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.41</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.36</td>
<td>0.55</td>
<td>0.53</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.52</td>
<td>0.40</td>
<td>0.25</td>
<td>0.44</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.44</td>
<td>0.26</td>
<td>0.10</td>
<td>-0.07</td>
<td>0.66</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.36</td>
<td>0.19</td>
<td>0.05</td>
<td>-0.17</td>
<td>0.42</td>
<td>0.87</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.30</td>
<td>0.19</td>
<td>0.07</td>
<td>-0.06</td>
<td>0.27</td>
<td>0.63</td>
<td>0.85</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.27</td>
<td>0.20</td>
<td>0.08</td>
<td>-0.02</td>
<td>0.31</td>
<td>0.51</td>
<td>0.65</td>
<td>0.84</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.14</td>
<td>0.07</td>
<td>0.01</td>
<td>-0.04</td>
<td>0.28</td>
<td>0.37</td>
<td>0.47</td>
<td>0.61</td>
<td>0.83</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.15</td>
<td>0.12</td>
<td>0.06</td>
<td>0.00</td>
<td>0.30</td>
<td>0.33</td>
<td>0.39</td>
<td>0.52</td>
<td>0.69</td>
<td>0.85</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.26</td>
<td>0.24</td>
<td>0.18</td>
<td>0.08</td>
<td>0.33</td>
<td>0.34</td>
<td>0.35</td>
<td>0.45</td>
<td>0.59</td>
<td>0.66</td>
<td>0.78</td>
<td>1.00</td>
</tr>
</tbody>
</table>

β denotes the month.
Table 3. Prediction scenarios and resulting efficiencies.

<table>
<thead>
<tr>
<th>Riv.</th>
<th>β</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>γ_m</th>
<th>γ_1</th>
<th>γ_0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amur</td>
<td>α</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>10</td>
<td>2</td>
<td>12</td>
<td>1</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.42</td>
<td>5.22</td>
<td>3.91</td>
<td>1.13</td>
<td>1.11</td>
<td>1.80</td>
<td>1.71</td>
<td>1.60</td>
<td>1.64</td>
<td>3.20</td>
<td>4.08</td>
<td>2.34</td>
<td>4.38</td>
<td>1.73</td>
<td>8.23</td>
</tr>
<tr>
<td>Irtysh</td>
<td>α</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>5.11</td>
<td>5.74</td>
<td>5.17</td>
<td>1.18</td>
<td>1.64</td>
<td>2.57</td>
<td>4.33</td>
<td>4.60</td>
<td>3.54</td>
<td>2.95</td>
<td>3.42</td>
<td>2.47</td>
<td>4.15</td>
<td>2.56</td>
<td>10.59</td>
</tr>
<tr>
<td>N.Dvi</td>
<td>α</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.61</td>
<td>1.36</td>
<td>1.02</td>
<td>1.26</td>
<td>1.08</td>
<td>1.15</td>
<td>2.04</td>
<td>1.67</td>
<td>1.48</td>
<td>0.93</td>
<td>1.95</td>
<td>1.48</td>
<td>2.95</td>
<td>1.21</td>
<td>3.58</td>
</tr>
<tr>
<td>Oka</td>
<td>α</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>10</td>
<td>36</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.59</td>
<td>1.29</td>
<td>1.25</td>
<td>0.96</td>
<td>1.92</td>
<td>1.30</td>
<td>1.33</td>
<td>1.90</td>
<td>1.72</td>
<td>1.40</td>
<td>2.66</td>
<td>3.01</td>
<td>1.39</td>
<td>1.40</td>
<td>2.02</td>
</tr>
<tr>
<td>Elba</td>
<td>α</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.29</td>
<td>1.17</td>
<td>1.04</td>
<td>1.38</td>
<td>1.28</td>
<td>1.21</td>
<td>1.21</td>
<td>1.27</td>
<td>1.43</td>
<td>1.27</td>
<td>1.56</td>
<td>1.72</td>
<td>1.27</td>
<td>1.27</td>
<td>1.61</td>
</tr>
<tr>
<td>Darm</td>
<td>α</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>12</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.31</td>
<td>1.21</td>
<td>1.08</td>
<td>1.46</td>
<td>1.78</td>
<td>1.47</td>
<td>1.14</td>
<td>1.52</td>
<td>2.02</td>
<td>1.45</td>
<td>1.78</td>
<td>2.16</td>
<td>1.09</td>
<td>1.41</td>
<td>1.57</td>
</tr>
<tr>
<td>Loire</td>
<td>α</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>12</td>
<td>11</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.21</td>
<td>1.15</td>
<td>1.07</td>
<td>1.21</td>
<td>1.45</td>
<td>1.93</td>
<td>1.76</td>
<td>1.52</td>
<td>1.41</td>
<td>1.17</td>
<td>1.32</td>
<td>1.81</td>
<td>1.47</td>
<td>1.31</td>
<td>2.00</td>
</tr>
<tr>
<td>Glom</td>
<td>α</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>2.62</td>
<td>5.31</td>
<td>2.97</td>
<td>1.06</td>
<td>1.10</td>
<td>0.97</td>
<td>1.33</td>
<td>1.22</td>
<td>1.32</td>
<td>1.26</td>
<td>1.74</td>
<td>1.39</td>
<td>3.39</td>
<td>1.21</td>
<td>3.90</td>
</tr>
</tbody>
</table>

Table 4. Validation of prediction scenarios for 5 rivers with longest intervals of observations.

<table>
<thead>
<tr>
<th>Riv.</th>
<th>β</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>γ_1</th>
<th>γ_0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irtysh</td>
<td>α</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>9</td>
<td>3</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>4.87</td>
<td>7.73</td>
<td>4.78</td>
<td>1.15</td>
<td>0.98</td>
<td>4.53</td>
<td>7.15</td>
<td>4.74</td>
<td>8.41</td>
<td>1.75</td>
<td>3.21</td>
<td>2.96</td>
<td>2.36</td>
<td>9.31</td>
</tr>
<tr>
<td>N.Dvi</td>
<td>α</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.39</td>
<td>4.86</td>
<td>1.01</td>
<td>1.00</td>
<td>0.99</td>
<td>1.46</td>
<td>1.02</td>
<td>1.64</td>
<td>1.55</td>
<td>1.56</td>
<td>1.00</td>
<td>1.23</td>
<td>1.12</td>
<td>6.80</td>
</tr>
<tr>
<td>Oka</td>
<td>α</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.65</td>
<td>1.65</td>
<td>1.01</td>
<td>1.05</td>
<td>0.91</td>
<td>1.22</td>
<td>1.94</td>
<td>1.77</td>
<td>1.08</td>
<td>1.66</td>
<td>3.77</td>
<td>1.19</td>
<td>1.06</td>
<td>3.02</td>
</tr>
<tr>
<td>Elba</td>
<td>α</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>12</td>
<td>1</td>
<td>10</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>2</td>
<td>2</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>1.68</td>
<td>1.36</td>
<td>1.18</td>
<td>1.44</td>
<td>1.46</td>
<td>1.78</td>
<td>0.87</td>
<td>1.11</td>
<td>1.15</td>
<td>1.34</td>
<td>1.52</td>
<td>1.67</td>
<td>1.36</td>
<td>1.61</td>
</tr>
<tr>
<td>Loire</td>
<td>α</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>11</td>
<td>11</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>10</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>0.92</td>
<td>1.19</td>
<td>1.27</td>
<td>1.22</td>
<td>1.29</td>
<td>2.10</td>
<td>4.26</td>
<td>1.65</td>
<td>1.10</td>
<td>1.08</td>
<td>1.04</td>
<td>1.64</td>
<td>1.23</td>
<td>2.02</td>
</tr>
</tbody>
</table>
Table 5. Values of $\kappa$, maximum possible number of common factors, and total efficiencies.

<table>
<thead>
<tr>
<th></th>
<th>Irtysh</th>
<th>Amur</th>
<th>N. Dvina</th>
<th>Oka</th>
<th>Glomma</th>
<th>Loire</th>
<th>Wisla</th>
<th>Elba</th>
<th>Danube</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_0$</td>
<td>10.59</td>
<td>8.23</td>
<td>6.65</td>
<td>3.58</td>
<td>3.90</td>
<td>2.00</td>
<td>2.02</td>
<td>1.61</td>
<td>1.57</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>2.56</td>
<td>1.75</td>
<td>1.27</td>
<td>1.21</td>
<td>1.21</td>
<td>1.31</td>
<td>1.40</td>
<td>1.27</td>
<td>1.41</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
Fig. 1. Monthly water runoffs of 9 rivers of Russia and Europe, N is number of readings in each series.
Fig. 2. The year forms (monthly average values) of runoff $m(\beta)$, $\beta=1,\ldots,12$ with standard deviation bars.
Fig. 3. Power spectra estimates for monthly water runoff.
Fig. 4. Deviations from the current estimates of the cyclic mean (signals $Y(t)$) for monthly water runoff.
Fig. 5. Sample estimates of variance in a moving time window (L=288, 24 years) for 4 rivers. The upper curve is the variance of initial time series $X(t)$ (see Fig. 1). The lower curves are different estimates of variance of stochastic component $Y(t)$ (see Fig. 4): dashed line is estimate of ANN-prediction error (eqn. (21)); the lowest curve is estimate of standard AR-prediction error in moving time window with AR-order $p=12$. All curves are given as functions of right-hand end of moving time window. Mark on time axis corresponds to the number of months since the beginning of observations.
Fig. 6. Deviations from the cyclic mean for Amur, months I-XII and graphs of 2 common orthogonal factors for deviations.