Forecasting of interest rate series

Prof. Dr. Ulrich Herkenrath
Gerhard-Mercator-Universität
Lotharstr. 65
47048 Duisburg
e-mail: herkenrath@math.uni-duisburg.de

Prof. Dr. Andreas Rudolph
Universität der Bundeswehr München
Fachbereich Betriebswirtschaft
85577 Neubiberg
e-mail: Andreas.Rudolph@UniBw-Muenchen.de

May 9, 2002

Contents

1 Introduction

In this paper we consider the problem of forecasting interest rates for bonds with M different maturities, e. g. with maturities m from 1 to 10 years. For a sequence of time periods, e. g. months, indexed by $t \in \mathbb{N}$, at a certain day in each period an interest rate is measured for each of the M maturities. If the measurement in time period t is present, then the forecasts for the M maturities should be given for all of the following H time periods, i. e. for each period $(t+h)$, $h = 1, \ldots, H$ and each maturity m, $1 \leq m \leq M$, the interest rate for maturity m, which will be valid in period $(t+h)$, should be forecasted on the basis of the measurements in periods 1 to t.

The observed values of interest rates in time periods 1 to t represent the "training set" for the forecasting procedure.

The practical problem of forecasting the interest rates for bonds with different maturities first leads to considerations how to choose an appropriate time series model. Although there is a large variety of models, it appears difficult to justify some concrete parametric model because of the consequences of such a choice.

A first question concerns the stationarity of the time series under consideration. Usually series of interest rates do not look stationary, but seem to contain a cyclic component. This in turn does not show a homogeneous periodicity concerning length and amplitudes. One may try to remove the cyclic component by a suitable "trend function" which is taken out of an admissible class of functions. The resulting series of residuals may be "nearer" to stationarity.

For the series of residuals one may think of the large class of ARIMA-models to conceive that series. After choosing such a model one is faced with the problem of constructing sequentially forecasts of step-lengths $1, \ldots, H$, $H = 12$.

In case that the h-step forecasts are computed by means of iterating the 1-step forecasts one finds quite empirically with respect to the "training set" of already observed values that with increasing h the errors of the forecasts increase drastically.

Therefore we propose a method which differs from the usual ones discussed in the literature (see e. g. Brockwell, Davis (1991), Lütkepohl (1991)). Our idea is to construct each h-step forecast, $1 \leq h \leq H$, independently on basis of the available training set by the method of conditional least squares, see Klimko, Nelson (1978).

From a formal viewpoint of course consistency problems arise because of the assumed recursive structure of the time series on one hand and the a-priori independence of the h-step forecast from the 1-step forecast or the $h-1$-step forecast in our method. We do not worry about that, but prefer our method, since we hope to grasp that way e. g. hidden periodic effects (the reason may be a complex dependence structure of the noise variables in the time series model) or to find eventually a simpler empirical h-step dependence within the series than may be derived from the assumed formal model.

Such a model is in general only a rough approximation of the real phenomenon.

In order to have a general solid basis for all those assumptions in a first part we present a theoretical framework with the results:

1. The sequence $(\boldsymbol{W}_t, t \geq 1)$ of interest rates for bonds with a fixed maturity represents a uniformly ergodic Markov-process.
2. As a consequence that process is asymptotically stationary, but in general not stationary just from the beginning. Uniform ergodicity of a homogeneous Markov process is equivalent to the mixing property of its 1-step transition probability, see e. g. Iosifescu, Grigorescu (1990).
3. That Markov process results from a reciprocal action of interest rates and exogeneous economic variables. The mathematical model to conceive the induced stochastic processes is a "generalized random system with complete connections" (abbreviated to GRSCC), see e. g. Iosifescu, Grigorescu (1990) . But we do not consider the classical huge and essential economic problem, to describe constructively and numerically that interdependence.
4. As a consequence we are faced with the problem to analyze the series of interest rates by detecting its internal structure. With regard to its Markovian structure we approximate the series of its residuals by a Markov process of order 3, e. g. by an ARMA(3,0)-process.

We shall propose the method of "conditional least squares" for the construction of the forecasts of the series. This method was introduced by Klimko, Nelson (1978) and is widely used in the meantime, see Brockwell, Davis (1991).

2 The theoretical model

We choose the mathematical framework of GRSCC's to model the evolution of interest rates in interaction with other economic variables. Under purely qualitative conditions on the interacting variables, like continuity and mixing properties of their stochastic interactions, we derive the uniform ergodicity of the process of interest rates.

For each time period $t \in \mathbb{N}$ we consider a p-dimensional vector $\boldsymbol{X}_t$, whose components represent economic parameters, like gross-national product etc. and a M-dimensional vector $\boldsymbol{W}_t$ of interest rates for the M different maturities. Those vectors are modelled as random vectors and their mutual dependencies by transition probabilities. Therefore let denote

- $W = [a,b]^M \subset \mathbb{R}^M$, $0 \leq a < b < \infty$, the range of possible interest rates for the M maturities,
- $X \subset \mathbb{R}^p$ the range of possible values of the p economic variables,
- $P : W \times \mathcal{X} \to [0,1]$ a transition probability from W to X, where $\mathcal{X}$ denotes the Borel-σ-algebra,
- $\pi : W \times X \times \mathcal{W} \to [0,1]$ a transition probability from $(W \times X)$ to W, where $\mathcal{W}$ also denotes the Borel-σ-algebra.

P represents the dependence of the economic variables from the preceding interest rates, i. e. if in period t $\boldsymbol{w}_t \in W$ is the vector of valid interest rates, then with probability $P(\boldsymbol{w}_t, B)$ the vector of economic variables will take on values in $B \in \mathcal{X}$ in period t. Analogously, if in period t $\boldsymbol{w}_t$ is the vector of interest rates and $\boldsymbol{x}_t$ the vector of economic parameters, then the random vector of interest rates in period $(t+1)$ is distributed according to the probability $\pi(\boldsymbol{w}_t, \boldsymbol{x}_t, .)$.

The theorem of Ionescu Tulcea ensures that for each $\boldsymbol{w} \in W$ as starting interest rate vector there exists a probability space $(\Omega, \mathcal{K}, P_{\boldsymbol{w}}) = ((W \times X)^{\mathbb{N}}, (\mathcal{W} \otimes \mathcal{X})^{\mathbb{N}}, P_{\boldsymbol{w}})$ and stochastic processes $(\boldsymbol{W}_t, t \geq 1)$ and $(\boldsymbol{X}_t, t \geq 1)$ with values in W respectively X such that

$$P_{\boldsymbol{w}}(\boldsymbol{W}_1 \in A) = \delta_{\boldsymbol{w}}(A)$$

$$P_{\boldsymbol{w}}(\boldsymbol{X}_1 \in B) = P(\boldsymbol{w}, B)$$

$$P_{\boldsymbol{w}}(\boldsymbol{W}_{t+1} \in A \mid \boldsymbol{W}_1, \boldsymbol{X}_1, \boldsymbol{W}_2, \boldsymbol{X}_2, \dots \boldsymbol{W}_t, \boldsymbol{X}_t) = \pi(\boldsymbol{W}_t, \boldsymbol{X}_t, A)$$

$$P_{\boldsymbol{w}}(\boldsymbol{X}_{t+1} \in B \mid \boldsymbol{W}_1, \boldsymbol{X}_1, \boldsymbol{W}_2, \boldsymbol{X}_2, \dots \boldsymbol{W}_t, \boldsymbol{X}_t, \boldsymbol{W}_{t+1}) = P(\boldsymbol{W}_{t+1}, B)$$

for $A \in \mathcal{W}$, $B \in \mathcal{X}$, where $\delta_{\boldsymbol{w}}$ denotes the one-point measure on $\{\boldsymbol{w}\}$. $E_{\boldsymbol{w}}$ will denote the mathematical expectation with respect to $P_{\boldsymbol{w}}$.

The above modelling of the economic variables , the interest rates and their mutual dependence in form of the transition probabilities means that the mathematical model which we choose represents a generalized random system with complete connections $\{W, X, \pi, P\}$ (abbreviated to GRSCC) in the sense of Iosifescu and Grigorescu (1990).

As a first important consequence one concludes: The process $(\boldsymbol{W}_t, t \geq 1)$ of interest rates constitutes a time-homogeneous Markov process with transition probability Q on $(W, \mathcal{W})$:

$$\begin{gathered} P_{\boldsymbol{w}}(\boldsymbol{W}_{t+1} \in A \mid \boldsymbol{W}_1, \boldsymbol{W}_2, \dots, \boldsymbol{W}_t) \\ = P_{\boldsymbol{w}}(\boldsymbol{W}_{t+1} \in A \mid \boldsymbol{W}_t) = \int_X \pi(\boldsymbol{W}_t, x, A) P(\boldsymbol{W}_t, dx) = Q(\boldsymbol{W}_t, A) \end{gathered}$$

In the terminology of the theory of the GRSCC's $(\boldsymbol{W}_t, t \geq 1)$ is the associated Markov process of the corresponding GRSCC. The Markov property enables to exploit the mighty theory on Markov processes, in particular ergodic theorems, to ensure desirable results on the process $(\boldsymbol{W}_t, t \geq 1)$, i. e. the evolution of interest rates.

The most important and desired property of a Markov process $(\boldsymbol{W}_t, t \geq 1)$ with transition probability Q on a state space $(W, \mathcal{W})$ is the so-called uniform ergodicity. That means:

There exists a probability measure π on $(W, \mathcal{W})$ such that

$$\lim_{n \to \infty} \sup_{\substack{\boldsymbol{w} \in W \\ A \in \mathcal{W}}} \mid Q^n(\boldsymbol{w}, A) - \pi(A) \mid = 0$$

where Q^n denotes the n-step transition probability of Q respectively $(\boldsymbol{W}_t, t \geq 1)$. Since for the Markov process $(\boldsymbol{W}_t, t \geq 1)$ with $\boldsymbol{W}_1 = \boldsymbol{w}$ the probability $Q^n(\boldsymbol{w}, .)$ represents the distribution of $\boldsymbol{W}_{n+1}$, uniform ergodicity amounts to a strong stability property of a Markov process. Another aspect of this property is the asymptotic stationarity of the process $(\boldsymbol{W}_t, t \geq 1)$, since the process with π as starting distribution would be a stationary process in the strong sense. Moreover the usual classical limit theorems hold true for uniformly ergodic Markov processes. We refer to the books of Iosifescu, Grigorescu (1990) and Meyn, Tweedie (1996) for a general more detailed discussion of uniform ergodicity.

We derive the desired property of the process $(\boldsymbol{W}_t, t \geq 1)$ essentially under continuity conditions of the transition probabilities P and π. Therefore we define the continuity property which is made use of.

Definition: Let $(Y, \mathcal{Y})$ and $(Z, \mathcal{Z})$ be measurable spaces, (Y, d) moreover a metric space and correspondingly $\mathcal{Y}$ die Borel-σ-algebra. A transition probability

S from $(Y, \mathcal{Y})$ to $(Z, \mathcal{Z})$ is called continuous if and only if for each $A \in \mathcal{Z}$ the mapping $S(., A) : (Y, d) \to [0, 1]$ is continuous.

A sufficient condition for the continuity of S is according to Scheffe's theorem, (see Herkenrath (1977)) : All measures $S(y, .)$ are dominated by a σ-finite measure λ on $(Z, \mathcal{Z})$ and for λ-almost all $z \in Z$ the densitites $s(., z)$ are continuous.

For this and related continuity concepts for transition probabilites see Herkenrath (1977) and Herkenrath (1979).

By means of well-known general results about associated Markov processes of GRSCC's under continuity assumptions we can show:

Theorem: Let the GRSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), \pi, P\}$ be as specified above, for each $\boldsymbol{x} \in X$ let $\pi_{\boldsymbol{x}}$ the transition probability on $(W, \mathcal{W})$, which is defined by $\pi_{\boldsymbol{x}}(\boldsymbol{w}, A) = \pi(\boldsymbol{w}, \boldsymbol{x}, A)$ for $\boldsymbol{w} \in W$, $A \in \mathcal{W}$.

In addition the transition probabilities P and π should satisfy:

(i) P is continuous,

(ii) π_x is continuous $\forall x \in X$,

(iii) there exists $B \in \mathcal{X}$ such that

$$\forall \boldsymbol{w} \in W \quad P(\boldsymbol{w}, B) > 0$$

and

$$\forall \boldsymbol{x} \in B, \boldsymbol{w} \in W \quad \pi(\boldsymbol{w}, \boldsymbol{x}, K_\epsilon) > 0$$

for all balls $K_\epsilon \subset W$ with positive radius ϵ.

Then the associated Markov process $(\boldsymbol{W}_t, t \geq 1)$ of the GRSCC is uniformly ergodic.

Proof: Due to Theorem 5.2 and Theorem 6.2 in Herkenrath (1979) the associated Markov process is "regular" with respect to $C(W)$, the Banach space of all complex valued bounded continuous functions on W, i. e. the corresponding Markov operator U is regular. According to Norman (1972, p. 40) under the assumptions on $(W, \mathcal{W})$ it follows regularity of U acting on $B(W)$, the Banach space of all complex valued bounded measurable functions on W. That in turn directly implies the uniform ergodicity of $(\boldsymbol{W}_t, t \geq 1)$, see e. g. Norman (1972, p. 38).

Remark: The continuity conditions on the transition probabilities do not represent heavy restrictions. A continuous variation of the external economic parameters from the interest rates seems plausible as well as a continuous variation of the new interest rates in dependence of the preceding ones for each fixed external situation. In assumption (iii) the set B of external parameters represents something like a set of "normal situations", which may be attained with positive probability from all preceding interest rates and which distributes the probability mass "over all W".

3 The Forecasting Procedure

The technique we propose to forecast the interest rates for the different maturities in the future is based on a quite heuristic approach.

Our studies lead us to the approach to deal each maturity separately, i. e. we do not forecast the vector $(W_{t+h}^{(1)}, \ldots, W_{t+h}^{(M)})$ for the interest rates of the M maturities in period $(t+h)$ on the basis of the vectors $(W_j^{(1)}, \ldots, W_j^{(M)})$, $1 \leq j \leq t$, but forecast $W_{t+h}^{(m)}$ on the basis of $W_j^{(m)}$, $1 \leq j \leq t$, for each $1 \leq m \leq M$.

So we take a fixed maturity m and its corresponding time series of interest rates, which we denote by $(\boldsymbol{W}_t, t \geq 1)$.

We do not assume a concrete model for the time series of interest rates which claims more than the Markov property and its ergodocity. As consequence the time series $(\boldsymbol{W}_t, t \geq 1)$ is asymptotically stationary. Since we want to forecast the future interest rates only by means of the past ones without taking into account the (external) econometric parameters, we base the forecast after the observation of $\boldsymbol{W}_T$ on the 3 last observations $\boldsymbol{W}_T, \boldsymbol{W}_{T-1}, \boldsymbol{W}_{T-2}$. In order to stabilize the forecasts we consider a "mean function" $m(.)$ in advance, so we forecast the residuals with respect to m instead of the values $\boldsymbol{W}_{T+h}$. The computation of the h-step forecasts is not calculated by a subsequent iteration of the 1-step forecasts, but is performed on the empirical basis of all a posteriori past h-step forecasts.

We propose this procedure in order to avoid an uncontrollable increase of the variance of the h-step forecasts caused by the iteration.

Our procedure works as follows:

1. Given the observations $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T)$ as training set we compute a mean function ${}_T m^{(m)} : \mathbb{N} \to W$ by minimizing the sum of squared deviations $\sum_{t=1}^{T} (W_t^{(m)} - f^{(m)}(t))^2$ for the m-th interest rate series within an admissible parametrized class of functions $f^{(m)} : \mathbb{R}_+ \to \mathbb{R}_+$, i. e.

$$\sum_{t=1}^{T}(W_t^{(m)} - {}_T m^{(m)}(t))^2 = \min_{f^{(m)}\in\mathcal{F}} \sum_{t=1}^{T}(W_t^{(m)} - f^{(m)}(t))^2, 1 \leq m \leq M$$

In case that a cyclic component appears within the time series $(W_t^{(m)}, t \geq 1)$, $\mathcal{F}$ may be chosen as the set of Fourier polynomials of a fixed degree r, e. g. 2, i. e.

$$\mathcal{F} = \{f \in \mathcal{F} \mid f(t) = a_0 + \sum_{i=1}^{r}[a_i \sin(2\pi\lambda_i t) + b_i \cos(2\pi\lambda_i t)]\}$$

with given frequencies $\lambda_i, i = 1, \ldots, r$. These frequencies will be chosen along the lines of Schlittgen and Streitberg (1999).

In case that a cyclic component is not clear at least within the training set, $\mathcal{F}$ may be a class of polynomials with a fixed degree, e. g. 1 or 2, or a parametrized class of exponential or logarithmic functions.

2. Next we remove the mean function and get the residuals $(R_t^{(m)}, 1 \leq t \leq T)$:

$$R_t^{(m)} = W_t^{(m)} - {}_T m^{(m)}(t)\,, 1 \leq t \leq T$$

Since we forecast the R_t's only by means of their own history, we model them as linear functions of their last 3 predecessors. We do that not step by step, but for each forecast step-length separately. Let H be the maximal step length. On basis of the training set $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T$ we define for the step-length h, $1 \leq h \leq H$, the estimator $\hat{R}_{T+h}^{(m)}$ for $R_{T+h}^{(m)}$ as:

$$\hat{R}_{T+h}^{(m)} = {}_T\alpha_h^{(m)} R_T^{(m)} + {}_T\beta_h^{(m)} R_{T-1}^{(m)} + {}_T\gamma_h^{(m)} R_{T-2}^{(m)},$$

where the parameters ${}_T\alpha_h^{(m)}$, ${}_T\beta_h^{(m)}$, ${}_T\gamma_h^{(m)}$, $1 \leq m \leq M$ are determined by

$$\sum_{t=3}^{T-h}(R_{t+h}^{(m)} - {}_T\alpha_h^{(m)} R_t^{(m)} - {}_T\beta_h^{(m)} R_{t-1}^{(m)} - {}_T\gamma_h^{(m)} R_{t-2}^{(m)})^2$$

$$= \min_{\alpha_h^{(m)},\beta_h^{(m)},\gamma_h^{(m)}\in\mathbb{R}} \sum_{t=3}^{T-h}(R_{t+h}^{(m)} - \alpha_h^{(m)} R_t^{(m)} - \beta_h^{(m)} R_{t-1}^{(m)} - \gamma_h^{(m)} R_{t-2}^{(m)})^2$$

for the fixed h.

The idea is that taking into account 3 preceding residuals, e. g. $R_t^{(m)}$, $R_{t-1}^{(m)}$, $R_{t-2}^{(m)}$, the sign of $R_{t+h}^{(m)}$ is estimated correctly.

3. On basis of the training set $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T$ after the computation of the H h-step forecasts $\hat{R}^{(m)}_{T+1}, \ldots, \hat{R}^{(m)}_{T+h}, \ldots, \hat{R}^{(m)}_{T+H}$, we gain the h-step forecasts $\hat{W}^{(m)}_{T+1}, \ldots, \hat{W}^{(m)}_{T+h}, \ldots, \hat{W}^{(m)}_{T+H}$ for the $W^{(m)}_{T+h}$-values, $1 \leq h \leq H$ for the m-th series, by incorporating the mean function values:

$$\hat{W}^{(m)}_{T+h} = \ {}_T m^{(m)}(T+h) + \hat{R}^{(m)}_{T+h}\,, 1 \leq h \leq H,\, 1 \leq m \leq M$$

4. After the observation of $\boldsymbol{W}_{T+1}$ the augmented training set $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_{T+1}$ is given, and the estimation procedure is performed again:

   In 1. the updated mean function ${}_{T+1}m^{(m)}$ is computed, in 2. the coefficients ${}_{(T+1)}\alpha^{(m)}_h$, ${}_{(T+1)}\beta^{(m)}_h$, ${}_{(T+1)}\gamma^{(m)}_h$ for the forecasts $\hat{R}^{(m)}_{(T+1)+h}$ of the H residuals $R^{(m)}_{(T+1)+1}, \ldots, R^{(m)}_{(T+1)+H}$ are determined and finally in 3. the forecasts $\hat{W}^{(m)}_{(T+1)+h}$, $1 \leq h \leq H$, $1 \leq m \leq M$ are obtained.

4 Discussion of the forecasting procedure

- Since with respect to the criterion of minimizing the mean squared error the forecasts $\hat{\boldsymbol{R}}_{T+h} = (\hat{R}^{(1)}_{T+h}, \ldots, \hat{R}^{(M)}_{T+h})$ should estimate $E_{\boldsymbol{w}}(\boldsymbol{R}_{T+h} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T)$ and the criterion for the construction of the estimators is that of "least squares", we denote our estimation procedure a method of "conditional least squares" (abbreviated to CLS). The notation refers to the work of Klimko and Nelson (1978), who studied the method of CLS for general ergodic processes.

- The conditional expectations obey a "consistency property" due to the fact that

$$E_{\boldsymbol{w}}(\boldsymbol{R}_{T+h} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T)$$
$$= \int E_{\boldsymbol{w}}(\boldsymbol{R}_{T+h} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T, \boldsymbol{R}_{T+1}) P_{\boldsymbol{w}}(d\boldsymbol{R}_{T+1} \mid \boldsymbol{R}_1, \ldots \boldsymbol{R}_T)\,, 1 < h \leq H$$

  In case that $\hat{\boldsymbol{R}}_{T+h}$ estimates $E_{\boldsymbol{w}}(\boldsymbol{R}_{T+h} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T)$, denoted by (*), then as a consequence a consistency property for the estimators respectively the coefficients ${}_T\alpha^{(m)}_h, {}_T\beta^{(m)}_h, {}_T\gamma^{(m)}_h$ and ${}_{(T+1)}\alpha^{(m)}_{h-1}, {}_{(T+1)}\beta^{(m)}_{h-1}$, ${}_{(T+1)}\gamma^{(m)}_{h-1}$ follows:

$$\hat{R}^{(m)}_{T+h} =$$
$${}_T\alpha^{(m)}_h R^{(m)}_T + {}_T\beta^{(m)}_h R^{(m)}_{T-1} + {}_T\gamma^{(m)}_h R^{(m)}_{T-2} =$$
$$\int \hat{R}^{(m)}_{T+1+(h-1)} P_{\boldsymbol{w}}(d\boldsymbol{R}_{T+1} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T) =$$

$$\int ({}_{(T+1)}\alpha_{h-1}^{(m)} R_{T+1}^{(m)} + {}_{(T+1)}\beta_{h-1}^{(m)} R_T^{(m)} + {}_{(T+1)}\gamma_{h-1}^{(m)} R_{T-1}^{(m)}) P_{\boldsymbol{w}}(d\boldsymbol{R}_{T+1} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T)$$

$$= {}_{(T+1)}\alpha_{h-1}^{(m)} E_{\boldsymbol{w}}(R_{T+1}^{(m)} \mid \boldsymbol{R}_1, \ldots, \boldsymbol{R}_T) + {}_{(T+1)}\beta_{h-1}^{(m)} R_T^{(m)} + {}_{(T+1)}\gamma_{h-1}^{(m)} R_{T-1}^{(m)}$$

$$= {}_{(T+1)}\alpha_{h-1}^{(m)} \hat{R}_{T+1}^{(m)} + {}_{(T+1)}\beta_{h-1}^{(m)} R_T^{(m)} + {}_{(T+1)}\gamma_{h-1}^{(m)} R_{T-1}^{(m)}$$

$$= ({}_{(T+1)}\alpha_{h-1}^{(m)} \, {}_T\alpha_1^{(m)} + {}_{(T+1)}\beta_{h-1}^{(m)}) R_T^{(m)}$$

$$+ ({}_{(T+1)}\alpha_{h-1}^{(m)} {}_T\beta_1^{(m)} + {}_{(T+1)}\gamma_{h-1}^{(m)}) R_{T-1}^{(m)}$$

$$+ {}_{(T+1)}\alpha_{h-1}^{(m)} {}_T\gamma_1^{(m)} R_{T-2}^{(m)}$$

The coupling between the coefficients $({}_{(T+1)}\alpha_{h-1}^{(m)}, {}_{(T+1)}\beta_{h-1}^{(m)}, {}_{(T+1)}\gamma_{h-1}^{(m)})$ and $({}_T\alpha_h^{(m)}, {}_T\beta_h^{(m)}, {}_T\gamma_h^{(m)})$ can of course recursively reduced to $h = 1$, so that the h-step estimators could be computed recursively from the 1-step forecasts.

Empirical studies show however that the iterated 1-step forecasts generate h-step forecasts which are worse than the h-step forecasts computed in one step according to our formula.

Therefore we do not take into consideration the consistency conditions on the coefficients $({}_T\alpha_h^{(m)}, {}_T\beta_h^{(m)}, {}_T\gamma_h^{(m)})$ for our CLS-estimation procedure.

- The h-step forecast corresponds to a h-step transition probability of the Markov process $(\boldsymbol{R}_t, t \geq 1)$. In case of the uniform ergodicity of that Markov process the sequence of the h-step transition probabilities converges to the same limit as the one of 1-step transition probabilities. Moreover, if $(R_t^{(m)}, t \geq 1)$ constitutes an AR(p)-process, for fixed $h \in \mathbb{N}$ the sequence $(R_{t+h}^{(m)}, t \geq 1)$ has an AR(p)-structure in the sense

$$R_{t+h}^{(m)} = \tilde{\alpha}_h^{(m)} R_t^{(m)} + \tilde{\beta}_h^{(m)} R_{t-1}^{(m)} + \tilde{\gamma}_h^{(m)} R_{t-2}^{(m)} + \tilde{\epsilon}_{t+h}^{(m)}$$

i. e. $({}_hY_t^{(m)} = R_{p+h-1+t}^{(m)}, t \geq 1)$ again represents an AR(p)-process.

Therefore the estimation of the parameters $({}_T\alpha_h^{(m)}, {}_T\beta_h^{(m)}, {}_T\gamma_h^{(m)})$ according to our procedure is reduced to the estimation of the "true parameters" for the AR(p)-process $({}_hR_t^{(m)}, t \geq 1)$.

Simultaneously the 1-step forecasts for $({}_hR_t, t \geq 1)$ are the h-step forecasts for $(R_t^{(m)}, t \geq 1)$. So for a fixed h, $1 \leq h \leq H$, concerning the $({}_hR_t^{(m)}, t \geq 1)$ series the convergence results for the estimators of the true parameters and the forecasts based on them , as they are demonstrated e. g. in Brockwell, Davis (1993) remain valid for the h-step forecasts for the series $(R_t^{(m)}, t \geq 1)$.

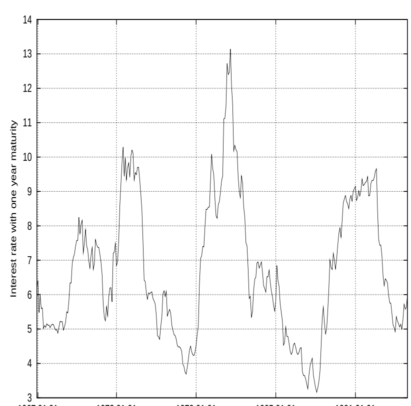
5 Numerical results

We have applied our procedure to the following three data sets:

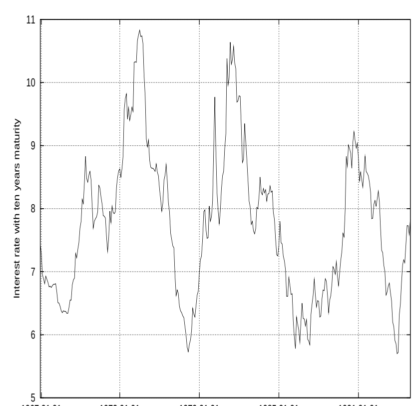
1. a data set containing the interest rates with one year maturity to ten year maturity, starting in January, 1967. The (for each series) 336 values are monthly, ending in December, 1994.
2. a data set containing the interest rates with one year maturity to ten year maturity, starting in January, 1994. The (for each series) 79 values are monthly too, ending in July, 2000.
3. a data set containing the interest rates with one year maturity to ten year maturity, starting in January, 1990. The (for each series) 129 values are monthly too, ending in September, 2000.

Series 2 and 3 are given at our disposal for testing purposes by the WestLB, whereas series 1 was due to a pension department of a large computer company.

To give some impression how these series look like, we consider here the interest rates with one respectively ten year maturity of series 1:

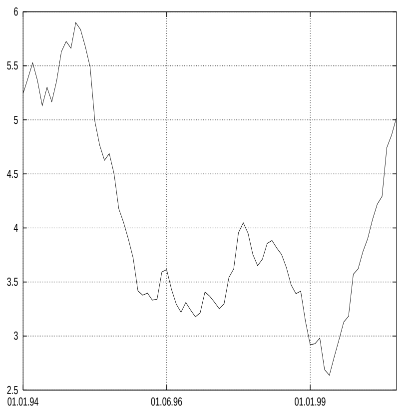


Interest rate with one year maturity

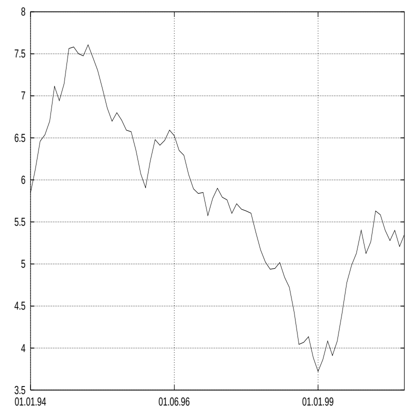


Interest rate with ten year maturity

and then the corresponding interest rates of series 2:

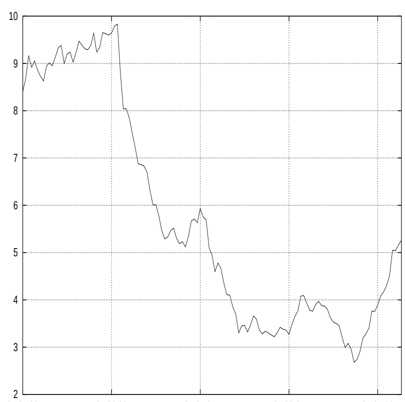


Interest rate with one year maturity

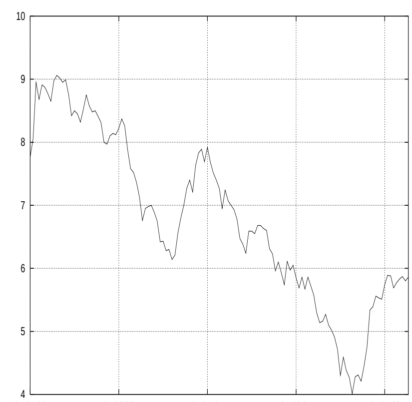


Interest rate with ten year maturity

Correspondingly the interest rates of Series 3 look like



Interest rate with one year maturity



Interest rate with ten year maturity

According to Filc (1992) univariate ARIMA-models should produce the best forecasts. He remarks that in the context of monthly interest rates the forecasts based on ARIMA-models seem to be optimal up to a forecast of step-length 3. It is not clear how his 3-step forecasts are calculated but one can suspect that they are iterated 1-step forecasts as recommended in the literature, see for example Box and Jenkins (1976), Brockwell and Davis (1993), Lütkepohl (1991) and Reinsel (1993), to cite only a few.

Now we want to compare our method with different ARIMA-models. To do so we use the following error measures to evaluate the forecast procedures on basis of an already known time series $(y_t, t = 1, \ldots, T + H)$, say.

For each fixed maturity m we get by our procedure for each time period $t = 1, \ldots, T$ and each forecast step h, $1 \leq h \leq H$ a forecast

$$\hat{y}_t(h) \text{ for the value } y_{t+h},$$

correspondingly an error

$$e_t(h) = y_{t+h} - \hat{y}_t(h)$$

To measure the performance of a forecasting procedure we consider for each maturity m and each forecast step length h (abbreviated to step h) the quantity

$$ {}^m\text{MAE(h)} = \frac{1}{T}\sum_{t=1}^{T} \mid y_{t+h} - \hat{y}_t(h) \mid \tag{1}$$

$$= \frac{1}{T}\sum_{t=1}^{T} \mid e_t(h) \mid$$

Moreover one could consider

$$ {}^m\text{MAX(h)} := \max_{1\leq t\leq T} \mid y_{t+h} - \hat{y}_t(h) \mid \tag{2}$$

and the following percentiles of the ordered sequence

$$\mid e_t(h) \mid_{(1)} \leq \ldots \leq \mid e_t(h) \mid_{(T)} :$$

$$ {}^m\text{QAE(h)} = \underbrace{(10\%, 25\%, 50\%, 75\%, 90\%)}_{\substack{\text{percentiles of the ordered} \\ \text{sequence } \mid e_t(h) \mid_{(1)} \leq \ldots \leq \mid e_t(h) \mid_{(T)}}} \tag{3}$$

To keep the amount of computed results within reasonable limits we will restrict ourselves on the measures ^{m}MAE(h) and ^{m}QAE(h) which we consider as impressive enough, although we could of course consider additional measures.

For a given data set which is evaluated by a forecasting procedure we get for the performance measure (1) a matrix, e. g.

$h\backslash m$	1	…	10	years of maturity
step 1	${}^{1}\mathrm{MAE}(1)$	…	${}^{10}\mathrm{MAE}(1)$	
step 2	${}^{1}\mathrm{MAE}(2)$		${}^{10}\mathrm{MAE}(2)$	
⋮	⋮		⋮	
step 12	${}^{1}\mathrm{MAE}(12)$	…	${}^{10}\mathrm{MAE}(12)$	
steps of forecasting				

One may think of calculating sums respectively means of rows or columns of this matrix. The mean of columns j may be denoted by

$$^{j}\overline{\mathrm{MAE}}(.),$$

the mean of rows i by

$$^{(.)}\overline{\mathrm{MAE}}(i).$$

With respect to measure (2) the column or row maxima may be calculated.

In the following we present the results.

For each of the three data sets which were mentioned above first we construct the forecasts according to an underlying ARIMA(2,1,0)-model, i. e. the series of interest rates for each maturity m is regarded as an ARIMA(2,1,0)-model and the h-step forecasts, $1 \leq h \leq H$, are computed as recommended in the literature, see for example Box, Jenkins (1976) or Brockwell, Davis (1993).

For an ARIMA(2,1,0)-model these are the optimal forecasts with respect to the criterion of minimizing the expected quadratic error. For $h > 1$ the forecasts are "iterated 1-step forecasts". We abbreviate this method as "the ARIMA(2,1,0)-model". Of course we tried additional ARIMA-models, but among those models the one with "parameters" (2,1,0) gave the best results.

Next we apply our method, which was described in Section 3, to each of the three data sets, separately for each maturity m too.

We used a Fourier polynomial of order 3 with the same data base as in the case of the ARIMA-model, but instead of using all available data points at a certain time point we took only a certain part of the xxx values (which means that we neglected more and more from the beginning of each series). The used xxx values were shown in the row with the label months. Please note that in case that the length of the considered series was smaller than xxx then of course we used of course the whole series.

For the calculation of the conditional least squares we took always the last 3 values.

The training set of data which gave the basis fore the forecasts is:

Since in any case we construct the forecasts "ex post", we can compare the forecasts with the real or true values for the corresponding interest rates. We measure the performance of both the ARIMA(2,1,0)-model and our method by means of the quantities (1) ^{m}MAE(h) and (3) ^{m}QAE(h).

We calculated the complete matrices of error measures (1) and (3) for each of the three data sets. The whole data sets of results is available under

http://www.bwi.unibw-muenchen.de/rudolph/index.html

Because of reasons of space here we present only a selection of these data, in particular we omit all results concerning data set 2. They are quite similar to these for data set 1 and 3.

5.1 Data set 1

We started our exploration with series 1 and calculated the error measures first for the above ARIMA-model and then for our approach.

5.1.1 Error measure MAE

The ARIMA(2,1,0)-model

years to maturity m	1	2	3	4	5
step h = step 1	0.41620	0.37818	0.35614	0.33998	0.32728
step 2	0.52590	0.46636	0.43542	0.41541	0.39915
step 3	0.60070	0.53459	0.49739	0.47447	0.45474
step 4	0.67859	0.60756	0.56618	0.53900	0.51542
step 5	0.74983	0.67072	0.62499	0.59618	0.57091
step 6	0.81198	0.72756	0.68213	0.64977	0.61942
step 7	0.88615	0.78289	0.73243	0.69717	0.66550
step 8	0.97533	0.84119	0.78100	0.74308	0.71216
step 9	1.08297	0.92135	0.84175	0.80400	0.77191
step 10	1.18961	1.00826	0.92275	0.87608	0.84000
step 11	1.30218	1.10259	1.01167	0.95941	0.91958
step 12	1.39850	1.19325	1.09475	1.03542	0.98833

^{m}MAE(h) for ARIMA(2,1,0) with maturities from 1 to 5 years

Our method

years to maturity m	1	2	3	4	5
months	215	210	210	205	203
step h = step 1	0.39413	0.34971	0.32578	0.31046	0.29618
step 2	0.47466	0.42463	0.39757	0.38166	0.36652
step 3	0.52759	0.47386	0.44743	0.42682	0.41602
step 4	0.55363	0.51352	0.49666	0.47876	0.46921
step 5	0.59059	0.54602	0.52070	0.52299	0.52417
step 6	0.61102	0.55920	0.54505	0.56032	0.56060
step 7	0.64132	0.57486	0.57407	0.59917	0.60254
step 8	0.67930	0.59911	0.60811	0.62525	0.63136
step 9	0.71730	0.63682	0.64613	0.66166	0.66774
step 10	0.74883	0.66693	0.68132	0.69439	0.70218
step 11	0.76942	0.69346	0.70479	0.71311	0.72247
step 12	0.78408	0.71539	0.73072	0.73539	0.74492

^{m}MAE(h) for our approach with maturities from 1 to 5 years

The ARIMA(2,1,0)-model

years to maturity m	6	7	8	9	10
step $h =$ step 1	0.31553	0.30494	0.29509	0.28665	0.27695
step 2	0.38407	0.37011	0.35713	0.34700	0.33724
step 3	0.43719	0.42108	0.40622	0.39405	0.38367
step 4	0.49828	0.48372	0.46913	0.45657	0.44607
step 5	0.54899	0.52983	0.51291	0.49850	0.48791
step 6	0.59275	0.57025	0.55091	0.53341	0.52016
step 7	0.63950	0.61733	0.59641	0.57758	0.56166
step 8	0.68850	0.66716	0.64650	0.62783	0.61058
step 9	0.74675	0.72566	0.70733	0.68933	0.67266
step 10	0.81183	0.78675	0.76508	0.74508	0.72741
step 11	0.88733	0.85958	0.83458	0.81233	0.79275
step 12	0.94991	0.91766	0.89058	0.86491	0.84150

^{m}MAE(h) for ARIMA(2,1,0) with maturities from 6 to 10 years

Our method

years to maturity m	6	7	8	9	10
months	202	202	202	202	202
step $h =$ step 1	0.28787	0.27906	0.27018	0.26206	0.25458
step 2	0.35175	0.34160	0.33478	0.32915	0.32335
step 3	0.40293	0.39496	0.38913	0.38300	0.37707
step 4	0.46048	0.45634	0.45234	0.44816	0.44295
step 5	0.51860	0.51344	0.50935	0.50392	0.49652
step 6	0.55564	0.55213	0.54663	0.54070	0.53331
step 7	0.59890	0.59413	0.58774	0.58030	0.57106
step 8	0.62986	0.62637	0.61873	0.61003	0.60055
step 9	0.66751	0.66420	0.65517	0.64536	0.63441
step 10	0.70027	0.69526	0.68585	0.67412	0.66199
step 11	0.72094	0.71650	0.70771	0.69823	0.68777
step 12	0.74761	0.74308	0.73546	0.72637	0.71543

^{m}MAE(h) for our approach with maturities from 6 to 10 years

5.1.2 Error measure QAE

Maturity $m = 1$ (short maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.06815	0.17085	0.31453	0.57262	0.92804
step 2	0.11791	0.20445	0.39041	0.75076	1.11374
step 3	0.05766	0.23483	0.45774	0.87609	1.30481
step 4	0.06836	0.22894	0.55744	0.94660	1.68679
step 5	0.05477	0.24211	0.65316	1.07296	1.78953
step 6	0.10935	0.23893	0.64207	1.28771	1.87931
step 7	0.11996	0.29926	0.74086	1.41981	2.02927
step 8	0.12014	0.32006	0.83974	1.56969	2.05772
step 9	0.15013	0.38991	0.94012	1.61009	2.25051
step 10	0.26003	0.46000	1.01004	1.70040	2.69013
step 11	0.27999	0.49005	1.11003	1.92005	2.88997
step 12	0.27008	0.61004	1.16003	2.04999	2.90000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 1$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.08148	0.15721	0.32206	0.55825	0.79576
step 2	0.09295	0.18814	0.36476	0.68413	1.08384
step 3	0.08239	0.18104	0.44376	0.75983	1.15635
step 4	0.08423	0.16647	0.45358	0.85300	1.23101
step 5	0.11331	0.24347	0.48116	0.83602	1.24987
step 6	0.09733	0.23392	0.50404	0.89401	1.34590
step 7	0.12615	0.23459	0.52977	1.02239	1.33033
step 8	0.14575	0.28321	0.55688	1.10252	1.36308
step 9	0.16225	0.30802	0.64849	1.05814	1.42103
step 10	0.17905	0.32117	0.69695	1.07958	1.56808
step 11	0.16581	0.35386	0.67330	1.10922	1.60600
step 12	0.17885	0.35154	0.72771	1.13748	1.61919

^{m}QAE(h) for our approach with maturity $m = 1$

Maturity $m = 5$ (mean maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.06117	0.13552	0.31180	0.48331	0.59421
step 2	0.08679	0.15599	0.33969	0.61313	0.77494
step 3	0.07926	0.16613	0.35939	0.72995	0.97488
step 4	0.08097	0.17157	0.38002	0.83049	1.09918
step 5	0.11993	0.22996	0.49976	0.77976	1.22011
step 6	0.11012	0.31000	0.51997	0.84000	1.29996
step 7	0.11997	0.27998	0.58001	0.92000	1.32000
step 8	0.09000	0.31000	0.69000	0.99001	1.46000
step 9	0.15000	0.27000	0.62000	1.19000	1.63000
step 10	0.15000	0.31000	0.62000	1.34000	1.85000
step 11	0.19000	0.36000	0.73000	1.53000	1.89000
step 12	0.16000	0.35000	0.90000	1.60000	1.95000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 5$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.07896	0.15522	0.27510	0.40224	0.54022
step 2	0.06615	0.17730	0.34070	0.52850	0.69218
step 3	0.05941	0.14386	0.39591	0.66421	0.79678
step 4	0.05694	0.16626	0.43892	0.68374	0.99040
step 5	0.09965	0.26786	0.50327	0.70597	1.01797
step 6	0.10800	0.25230	0.54277	0.79909	1.04088
step 7	0.11625	0.29831	0.57740	0.88524	1.10202
step 8	0.07828	0.29334	0.62077	0.98919	1.20116
step 9	0.08747	0.28003	0.60669	0.99788	1.26321
step 10	0.12653	0.29668	0.64314	1.04202	1.39252
step 11	0.10722	0.27906	0.71584	1.12571	1.47000
step 12	0.08406	0.25697	0.70221	1.19665	1.53343

^{m}QAE(h) for our approach with maturity $m = 5$

Maturity $m = 10$ (long maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.05444	0.12237	0.23820	0.42080	0.54941
step 2	0.03963	0.14877	0.29973	0.52216	0.66750
step 3	0.04579	0.14013	0.32887	0.61970	0.82986
step 4	0.07009	0.16021	0.38001	0.69988	0.98000
step 5	0.05999	0.16000	0.41002	0.65000	1.05001
step 6	0.06000	0.20000	0.48000	0.76000	1.09000
step 7	0.08000	0.22000	0.47000	0.81000	1.07000
step 8	0.09000	0.29000	0.51000	0.84000	1.29000
step 9	0.18000	0.30000	0.54000	0.94000	1.44000
step 10	0.19000	0.36000	0.62000	1.05000	1.48000
step 11	0.18000	0.39000	0.66000	1.18000	1.63000
step 12	0.14000	0.37000	0.70000	1.25000	1.71000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 10$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.03830	0.10404	0.23254	0.35958	0.53961
step 2	0.07161	0.14396	0.27128	0.50003	0.64204
step 3	0.08448	0.15343	0.35351	0.60900	0.74949
step 4	0.11943	0.21064	0.39593	0.64575	0.86373
step 5	0.14955	0.24807	0.49293	0.72024	0.91882
step 6	0.15828	0.27627	0.55799	0.79933	0.94300
step 7	0.15479	0.30322	0.55739	0.83004	1.04099
step 8	0.12291	0.33755	0.58805	0.90228	1.06977
step 9	0.14334	0.38342	0.60915	0.94012	1.10286
step 10	0.12822	0.36542	0.65843	0.93471	1.16672
step 11	0.12134	0.32092	0.69025	0.95017	1.20618
step 12	0.10364	0.40369	0.74247	1.01458	1.28669

^{m}QAE(h) for our approach with maturity $m = 10$

5.2 Data set 3

The data base consisted of the first 40 months, with these months 76 forecasts for the next 12 months were calculated sucessively. In these two comparisons we took all data points.

The Fourier polynomial consisted here of a polynomial of degree 1, for the forecasts with the conditional least squares we took the last three residuals.

5.2.1 Error measure MAE

The ARIMA(2,1,0)-model

years to maturity m	1	2	3	4	5
step $h =$ step 1	0.25455	0.31033	0.33119	0.33259	0.32414
step 2	0.32037	0.39259	0.42735	0.42795	0.41304
step 3	0.39478	0.48530	0.52625	0.53276	0.52051
step 4	0.46089	0.57824	0.62224	0.62169	0.60394
step 5	0.52456	0.65654	0.70904	0.71252	0.69354
step 6	0.58015	0.73050	0.77818	0.78407	0.76277
step 7	0.63394	0.77430	0.82750	0.82710	0.80224
step 8	0.68552	0.81762	0.86592	0.87236	0.84855
step 9	0.73052	0.86776	0.90947	0.90592	0.88763
step 10	0.77973	0.90763	0.94131	0.93631	0.91394
step 11	0.82421	0.94907	0.96907	0.95552	0.93578
step 12	0.86763	0.97881	0.98934	0.96750	0.95078

^{m}MAE(h) for the ARIMA-model with maturities from 1 to 5 years

Our method

years to maturity m	1	2	3	4	5
step $h =$ step 1	0.23856	0.29856	0.31456	0.31574	0.30513
step 2	0.31132	0.38401	0.41038	0.40917	0.39266
step 3	0.37313	0.46386	0.49580	0.49485	0.47818
step 4	0.42859	0.52906	0.56680	0.56546	0.54735
step 5	0.48126	0.60926	0.65314	0.64714	0.62231
step 6	0.52328	0.65500	0.69927	0.69495	0.66571
step 7	0.56351	0.68380	0.72265	0.72235	0.69586
step 8	0.59771	0.70574	0.73676	0.73299	0.71053
step 9	0.61491	0.72052	0.73429	0.73216	0.71237
step 10	0.64191	0.74286	0.75937	0.74017	0.71378
step 11	0.66016	0.74949	0.76406	0.76082	0.73272
step 12	0.68126	0.75758	0.77923	0.77852	0.76010

^{m}MAE(h) for our approach with maturities from 1 to 5 years

The ARIMA(2,1,0)-model

years to maturity m	6	7	8	9	10
step h = step 1	0.31257	0.29569	0.27677	0.25740	0.24638
step 2	0.39685	0.38258	0.36322	0.34321	0.32988
step 3	0.50532	0.48398	0.45789	0.43318	0.41461
step 4	0.58352	0.55672	0.52763	0.50275	0.48472
step 5	0.66964	0.63972	0.61191	0.58479	0.56738
step 6	0.74015	0.70912	0.68058	0.64794	0.62925
step 7	0.78184	0.75857	0.73316	0.70368	0.68289
step 8	0.82841	0.80500	0.77579	0.74487	0.72434
step 9	0.86381	0.83789	0.80855	0.78210	0.76526
step 10	0.89565	0.87789	0.85184	0.82447	0.80657
step 11	0.92263	0.91052	0.88750	0.86407	0.84776
step 12	0.94210	0.94013	0.92315	0.89881	0.88276

^{m}MAE(h) for the ARIMA-model with maturities from 6 to 10 years

Our method

years to maturity m	6	7	8	9	10
step h = step 1	0.29391	0.27642	0.25882	0.23898	0.22678
step 2	0.37714	0.35717	0.33485	0.31255	0.29923
step 3	0.46217	0.44190	0.41518	0.39162	0.37474
step 4	0.52678	0.50165	0.47104	0.44586	0.42681
step 5	0.59418	0.56368	0.53172	0.50330	0.48006
step 6	0.63909	0.60890	0.57786	0.54830	0.52445
step 7	0.67011	0.64302	0.61223	0.58142	0.55931
step 8	0.69369	0.67151	0.64091	0.61175	0.59306
step 9	0.69861	0.67899	0.65694	0.63241	0.61520
step 10	0.70218	0.69444	0.67547	0.65748	0.64841
step 11	0.72011	0.70835	0.69382	0.67531	0.66695
step 12	0.74152	0.72438	0.70425	0.68393	0.67946

^{m}MAE(h) for our approach with maturities from 6 to 10 years

5.2.2 Error measure QAE

Maturity $m = 1$ (short maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.05089	0.09908	0.21118	0.36393	0.51500
step 2	0.02917	0.15921	0.28735	0.49187	0.66198
step 3	0.05107	0.14013	0.36982	0.60208	0.81864
step 4	0.07947	0.13967	0.40718	0.72088	0.96691
step 5	0.06002	0.19991	0.44007	0.80997	1.08014
step 6	0.10000	0.21024	0.52014	0.86077	1.20999
step 7	0.08000	0.26000	0.55000	0.94001	1.32999
step 8	0.14000	0.28000	0.51000	1.01988	1.40000
step 9	0.19000	0.33000	0.56002	1.12998	1.51999
step 10	0.15000	0.34000	0.63000	1.20000	1.59999
step 11	0.13000	0.39000	0.67000	1.14000	1.64000
step 12	0.15000	0.41000	0.70000	1.23000	1.86000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 1$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.04117	0.10919	0.20363	0.34343	0.48923
step 2	0.06092	0.14725	0.29810	0.43641	0.62913
step 3	0.03503	0.15244	0.37923	0.60523	0.70589
step 4	0.05607	0.15663	0.43434	0.65539	0.84026
step 5	0.08509	0.18390	0.45194	0.76143	0.97747
step 6	0.08018	0.24587	0.44342	0.79576	1.10675
step 7	0.12577	0.24806	0.47619	0.92920	1.16799
step 8	0.10886	0.29968	0.51358	0.82647	1.24470
step 9	0.05232	0.25466	0.59723	0.85175	1.27095
step 10	0.11753	0.25635	0.59966	0.92495	1.29679
step 11	0.09212	0.26663	0.59608	0.96953	1.31257
step 12	0.11865	0.23293	0.63156	1.05155	1.38006

^{m}QAE(h) for our approach with maturity $m = 1$

Maturity $m = 5$ (mean maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step $h =$ step 1	0.07162	0.14186	0.27397	0.47117	0.63601
step 2	0.06100	0.18369	0.36869	0.64791	0.83715
step 3	0.11006	0.23029	0.49586	0.79164	1.02717
step 4	0.06113	0.28827	0.57035	0.90019	1.19963
step 5	0.17076	0.31985	0.72007	1.07027	1.29978
step 6	0.16012	0.33962	0.73994	1.17008	1.39000
step 7	0.09006	0.29001	0.81996	1.28000	1.43991
step 8	0.11999	0.42999	0.85001	1.20998	1.64997
step 9	0.20000	0.40000	0.83998	1.33000	1.80000
step 10	0.15000	0.47000	0.82000	1.29000	1.83000
step 11	0.19000	0.40000	0.84000	1.43000	1.97000
step 12	0.08000	0.42000	0.94000	1.54000	1.95000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 5$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step $h =$ step 1	0.05713	0.13657	0.28536	0.42583	0.55510
step 2	0.04221	0.18147	0.39261	0.56991	0.74201
step 3	0.06177	0.23664	0.49757	0.72726	0.91993
step 4	0.07049	0.24111	0.46170	0.81991	1.07328
step 5	0.12509	0.28241	0.57944	0.92026	1.10551
step 6	0.09298	0.28710	0.61064	0.90144	1.37480
step 7	0.12016	0.32439	0.61804	0.91784	1.46234
step 8	0.13238	0.32155	0.50393	1.02386	1.55790
step 9	0.10039	0.25975	0.52937	1.08705	1.49697
step 10	0.04190	0.14541	0.52801	1.22817	1.53045
step 11	0.07116	0.23309	0.47119	1.14725	1.68603
step 12	0.17466	0.27257	0.53206	1.06399	1.61035

^{m}QAE(h) for our approach with maturity $m = 5$

Maturity $m = 10$ (long maturity)

The ARIMA(2,1,0)-model

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.04448	0.09068	0.23609	0.37561	0.47096
step 2	0.06882	0.12763	0.28999	0.50797	0.65367
step 3	0.09201	0.20000	0.41069	0.56411	0.78327
step 4	0.07721	0.23237	0.47180	0.70001	0.96855
step 5	0.14000	0.34010	0.52000	0.75000	1.03002
step 6	0.19000	0.36952	0.58005	0.86000	1.14040
step 7	0.15994	0.38998	0.65000	0.91999	1.20006
step 8	0.10005	0.44988	0.70002	0.99000	1.28998
step 9	0.18000	0.39000	0.76000	1.13000	1.41999
step 10	0.24000	0.45000	0.79000	1.13000	1.55000
step 11	0.27000	0.51000	0.83000	1.26000	1.47000
step 12	0.32000	0.50000	0.84000	1.31000	1.54000

^{m}QAE(h) for ARIMA(2,1,0) with maturity $m = 10$

Our method

fractiles	0.1	0.25	0.5	0.75	0.9
step h = step 1	0.02660	0.06155	0.19766	0.33681	0.48056
step 2	0.05453	0.11125	0.25079	0.41578	0.67860
step 3	0.02785	0.19217	0.33839	0.52851	0.71816
step 4	0.07234	0.17051	0.36818	0.60366	0.87961
step 5	0.06120	0.18384	0.43113	0.72217	1.07940
step 6	0.08811	0.17992	0.47282	0.68660	1.06329
step 7	0.05693	0.21015	0.42712	0.79211	1.29907
step 8	0.11415	0.20089	0.43416	0.83119	1.48525
step 9	0.06682	0.23470	0.46230	0.92161	1.42873
step 10	0.11238	0.20270	0.46413	1.03735	1.42163
step 11	0.07771	0.20017	0.49901	0.99971	1.35094
step 12	0.05478	0.18373	0.49460	1.02165	1.56608

^{m}QAE(h) for our approach with maturity $m = 10$

6 Discussion of the results

As one can see, the accuracy of our method is ”nearly uniformly” better then the traditional technique based on the ARIMA-models. For a lot of cases, e. g.

large values of h and m, we get a drastic improvement. The term "uniformly" refers to the different data sets, the two error measures and to the pairs (m, h).

For the mean absolute error it can be seen for example in case of the interest series 1 with one year maturity that one gets an error of 1.3985 for a twelve months forecast, whereas for the same series in case of our approach we have for a twelve months forecast an error of 0.78408, which means in this case a dramatic improvement. This is quite remarkable because usually interest series with one year maturity are considered to be much more difficult to forecast than series with longer maturities.

If we compare the case of series 1 with 10 year maturity we have for the twelve months forecast a mean absolute error of 0.84159 for the ARIMA-model whereas in case of our approach we get an mean absolute error of 0.71543. Similar results can be found also in case of the other series.

In case of the QAE(h) measure it can be seen that the length of the intervals given by for example the 0.9 quantile and the 0.1 quantile is shorter than the corresponding one of the ARIMA-models and also nearer to zero, which means a better precision than the counterpart given by the ARIMA-models, too. As an example take the interest series 1 with one year maturity and a twelve months forecast for the 0.9 fractile which gives 2.9 and the 0.1 fractile which gives 0.27008 in case of the ARIMA-model whereas in our approach we get for the same series for the 0.9 fractile a value of 1.61919 and for the 0.1 fractile a value of 0.17885 which means in case of the ARIMA-model an interval of length 2.62992 and in case of our approach a length of 1.44034 which means a much greater precision in case of our approach.

Literature

Box, G. E. P., and Jenkins, G. M. (1976). Time Series Analysis: forecasting and control, Holden-Day

Brockwell, P. J. and Davis, R. A. (1993). Time Series: Theory and Methods. (Second edition), Springer, New York

Filc, W. (1992). Theorie und Empirie des Kapitalmarktzinses, Deutscher Sparkassenverlag GmbH, Stuttgart

Herkenrath, U. (1977). Markov processes under continuity assumptions, Revue Roumaine Math. Pures et Appl. 22, 1419 - 1432

Herkenrath, U. (1979). The associated Markov process of generalized random systems with complete connections, Revue Roumaine Math. Pures et Appl. 24, 243 - 254

Iosifescu, M. and Grigorescu, S. (1990). Dependence with Complete Connections and its Applications, CUP, Cambridge

Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes, Annals of Statistics 6, 629 - 642

Meyn, S. P. and Tweedie, R. L. (1996). Markov Chains and Stochastic Stability, Springer, London

Lütkepohl, H.(1991). Introduction to Multiple Time Series Analysis, Springer, New York

Norman, M. F. (1972). Markov Processes and Learning Models, Academic Press, New York

Reinsel, G. C.(1993). Elements of Multivariate Time Series, Springer, New York

Schlittgen R. and Streitberg. B. H. J. (1999). Zeitreihenanalyse (8-th edition). Oldenbourg, Munich