Proceedings of ALGORITMY 2020
pp. 211–220

CALIBRATION OF THE VASICEK MODEL OF INTEREST RATES USING BICRITERIA OPTIMIZATION*

TATIANA JAŠURKOVÁ† AND BEÁTA STEHLÍKOVÁ‡

Abstract. The Vasicek model of the interest rates is one of the most frequently used short rate models to describe the movements of the interest rates. For the model to work properly it has to be adequately calibrated. Based on different approaches, there are several techniques to calibrate the Vasicek model. In this paper, we combine two criteria: fitting term structures of the interest rates and comparison of the estimated short rate with its estimate from the Kalman filter, which takes probability distributions into account. Doing so, we obtain the risk-neutral parameters as well as the estimate for the short rate. The proposed algorithm is then applied to the real market data and we analyze the results.

Key words. Vasicek model, calibration, Kalman filter, term structure, bicriteria optimization

AMS subject classifications. 91G30, 47N10, 90C29, 93E11

1. Introduction. The term structure of the interest rates describes the relationship between the maturities and the respective interest rates $R(t,T)$ of a set of discount bonds and is denoted by $P(t,T) = e^{-R(t,T)(T-t)}$, where $P(t,T)$ is the price of the above-mentioned discount bond.

When modelling the term structure of the interest rates, one of the popular choices is to use short rate models (other possibilities include LIBOR market models, models fitting the initial term structure, etc.; we refer the reader to [5], [15] or [19] for a detailed treatment of interest rate modelling. Short rate models assume that the short rate r_t, i.e., instantaneous interest rate, follows a specific stochastic process defined by a stochastic differential equation. In these models, the price of the discount bond is a function of the time, its maturity and the current level of the short rate. After the specification of so-called market price of risk the bond price then satisfies a parabolic partial differential equation. In this paper we consider the Vasicek model [23] which models the short rate by an Ornstein-Uhlenbeck process which results in the normal distribution. Furthermore, for a specific market price of risk the bond prices are known in a closed form. There are many different ways for the calibration of this model. In this paper, we combine two of them, one of which uses the probability distributions (Kalman filtering) and one which compares real and theoretical term structures of interest rates.

The paper is organized as follows. In Section 2 we shortly revise the Vasicek model and state its properties which we need in subsequent calibration. Section 3 provides a brief overview of calibration methods and describes two of them in a detail, which will serve as a base of our approach. In Section 4 we explain the basic concepts of multicriteria optimization problems. Our main contribution is contained in Section 5 where we describe our calibration procedure and in Section 6 where we apply it to real market data. We end the paper with concluding remarks in Section 7.

*This work was supported by VEGA 1/0062/2018 grant and Grant UK/29/2020.
†Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia (jasurkova.tatiana@gmail.com).
‡Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 842 48 Bratislava, Slovakia (stehlikova@fmph.uniba.sk).

2. Vasicek model of interest rates. Vasicek model [23] belongs to the class of one-factor short rate models. In particular, the dynamics of the short rate is governed by the Ornstein-Uhlenbeck process

$$dr_t = \kappa(\theta - r_t)dt + \sigma dw_t, \tag{2.1}$$

where κ, θ, σ are model parameters and w_t is a Wiener process. Subsequently, short rate is then normally distributed (cf. [23])

$$r_{t+\triangle t} \mid r_t \sim \mathcal{N}\left(\theta\left(1 - e^{-\kappa \triangle t}\right) + e^{-\kappa\triangle t} r_t, \quad \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\triangle t}\right)\right). \tag{2.2}$$

In the Vasicek model market price of risk is typically chosen to be a constant, i.e., $\lambda(r,t) = \lambda$. Then, the partial differential equation for the bond price $P = P(r,\tau)$, where $\tau = T - t$ is time remaining to maturity of the bond, satisfies the partial differential equation (cf. again [23])

$$-\frac{\partial P}{\partial \tau} + (\kappa(\theta - r) - \lambda\sigma)\frac{\partial P}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 P}{\partial r^2} - rP = 0 \tag{2.3}$$

with an initial condition $P(r,0) = 1$. The solution to this equation can be found in an explicit form; we write it in the form from [19], since it will be useful in subsequent computations

$$P(r,\tau) = A(\tau)e^{-B(\tau)r}, \tag{2.4}$$

$$\ln A(\tau) = \left[\frac{1}{\kappa}\left(1 - e^{-\kappa\tau}\right) - \tau\right]\left(\theta - \frac{\lambda\sigma}{\kappa} - \frac{\sigma^2}{2\kappa^2}\right) - \frac{\sigma^2}{4\kappa^3}(1 - e^{-\kappa\tau})^2, \tag{2.5}$$

$$B(\tau) = \frac{1}{\kappa}\left(1 - e^{-\kappa\tau}\right). \tag{2.6}$$

3. Various methods of calibration short rate models. There are many different approaches to calibrating short rate models. One of them is to consider probability distribution of the short rate and apply maximum likelihood estimation or its approximations (see, for example, [1], [20]), generalized method of moments ([10]), etc. Probability distributions, but those of interest rates with other maturities, are used when applying Kalman filter, such as in [3], [4] and [14], which leads also to the estimation of the underlying short rate. Fitting term structures of interest rates was used in [22]; the likelihood function was used only in the second step when choosing among parameter sets leading to the same fit. In the opposite way, [11] uses short rate distribution to estimate the real parameters and term structures of interest rates to estimate the market prices of risk. Papers [8] and [9] consider only fit of the term structures and hence estimate the risk-neutral parameters, they also estimate the short rate based on this criterion.

In the following two subsections we review the two approaches which form the base of our proposed algorithm. Both take the short rate as an unknown quantity and estimate it from the term structures, but use different ways to achieve this goal.

3.1. Kalman filter. Kalman filter, originally introduced by Rudolph E. Kalman in 1960 (see [16]), is a recursive algorithm used to filter out the true signal from a stream of measured data that is affected by a noise. Generally, these observed values also depend on other latent state variables. In our case, the set of measured data consists of the observations of bond rates and the state variable is identified with the realizations of the short rate. For more information about the Kalman filter we refer to [7], [18], or [21]. In this subsection the Kalman filter applied on the Vasicek model is described based on the work done in [4].

In order to use Kalman filtering as a method of estimating the parameters of the Vasicek model, it is necessary to rewrite the model into the state-space form, which involves the specification of the measurement system and the transition system. Firstly let consider a time window that consists of n days, further denoted as $t_1, t_2, \dots t_n$. We also assume that in each time step t_i, $i = 1, \dots, n$, there are m zero-coupon bonds, each one with a different length of maturity $\tau_1, \tau_2, \dots, \tau_m$, traded at the market. The measurement system then consists of the equations for the observed variable, i.e., bond rates $R(r_i, \tau_j)$. According to the solution of discount bond in the Vasicek model (2.4)-(2.6) the measurement system has the following form

$$\begin{bmatrix} R(r_i,\tau_1) \\ R(r_i,\tau_2) \\ \vdots \\ R(r_i,\tau_m) \end{bmatrix} = \begin{bmatrix} -\frac{\ln A(\tau_1)}{\tau_1} \\ -\frac{\ln A(\tau_2)}{\tau_2} \\ \vdots \\ -\frac{\ln A(\tau_m)}{\tau_m} \end{bmatrix} + \begin{bmatrix} \frac{B(\tau_1)}{\tau_1} \\ \frac{B(\tau_2)}{\tau_2} \\ \vdots \\ \frac{B(\tau_m)}{\tau_m} \end{bmatrix} \begin{bmatrix} r_{t_i} \end{bmatrix} + \begin{bmatrix} v_1(t_i) \\ v_2(t_i) \\ \vdots \\ v_m(t_i) \end{bmatrix} \tag{3.1}$$

or shortly $R_{t_i} = A + Hr_{t_i} + v(t_i)$, where $v(t_i)$ represents the measurement noise, which is assumed to be normally distributed with zero mean and covariance matrix R. When applying Kalman filter on interest rates, measurement noise may account for the data-entry errors, bid-ask spreads, or non-simultaneous observations. Under the assumption that the components of the $v(t_i)$ are not correlated the covariance matrix R can be written in the form of positive definite diagonal matrix, i.e., $R = diag(s_1^2, s_2^2, \dots, s_m^2)$.

On the other hand, the transition system describes the dynamics of the latent factor, i.e., realizations of the short rate. As in the Vasicek model short rate follows the stochastic process (2.1), short rate is therefore normally distributed (2.2) and the transition equation is of the form

$$r_{t_i} = \theta\left(1 - e^{-\kappa \triangle t}\right) + e^{-\kappa \triangle t} r_{t_{i-1}} + \varepsilon(t_i), \text{ or } r_{t_i} = C + F r_{t_{i-1}} + \varepsilon(t_i), \tag{3.2}$$

where $\triangle t = 1/252$, as daily observations are considered, and process noise $\varepsilon(t_i)$ is normally distributed, i.e., $\varepsilon(t_i) \mid r_{t_{i-1}} \sim \mathcal{N}\left(\quad 0, \quad \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa \triangle t}\right)\right)$. Thus the state-space form of the Vasicek one-factor model is then represented by the equations (3.1) and (3.2). The Kalman filter applied to this state-space form then makes an educated guess about the initial value of the unobserved state variable and predicts the value of the measurement equations. Subsequently, based on the error between the measurement equation prediction and the actual realization, the value of the state variable is updated and used in order to predict the next state. After recursing through all of the time steps, a time series for the short rate, or the filtration $\mathcal{F}_s$, is obtained. More precisely $\mathcal{F}_{t_i} = \sigma(R_{t_0}, R_{t_1}, \dots, R_{t_i})$, for $i = 1, \dots, n$, stands for the filtration of the short rate based on the first i days observations of the interest rate of m zero-coupon yields. In our case, the Kalman filter proceeds as follows:

Step 1. The unconditional mean and variance of the state variable are taken as the initial values of the state variable and the measure of the certainty of this guess, i.e.,

$$E[r_{t_1}] = E[r_t \mid \mathcal{F}_0] = \theta, \quad Var[r_{t_1}] = Var[r_{t_1} \mid \mathcal{F}_0] = \frac{\sigma^2}{2\kappa}.$$

Step 2. The value of the measurement system, i.e., the rates R_{t_i}, are forecasted as

$$E[R_{t_i} \mid \mathcal{F}_{t_{i-1}}] = A + HE[r_{t_i} \mid \mathcal{F}_{t_{i-1}}], \quad Var[R_{t_i} \mid \mathcal{F}_{t_{i-1}}] = HVar[r_{t_i} \mid \mathcal{F}_{t_{i-1}}]H^T + R.$$

Step 3. To improve the estimation of the state variable, filter compares the estimation and the realized value of the output variable, which generates a vector of prediction errors $\zeta_{t_i} = R_{t_i} - E[R_{t_i} \mid \mathcal{F}_{t_{i-1}}]$. This step also includes the update of the state variable and its conditional variance given by the following formulae

$$E[r_{t_i} \mid \mathcal{F}_{t_i}] = E[r_{t_i} \mid \mathcal{F}_{t_{i-1}}] + K_{t_i}\zeta_{t_i}, \quad Var[r_{t_i} \mid \mathcal{F}_{t_i}] = (I - K_{t_i}H)Var[R_{t_i} \mid \mathcal{F}_{t_{i-1}}],$$

where $K_{t_i} = Var[r_{t_i} \mid \mathcal{F}_{t_{i-1}}]H^T Var[R_{t_i} \mid \mathcal{F}_{t_{i-1}}]^{-1}$ is Kalman gain matrix.
Step 4. The last step consists of predicting the values of the state variable

$$E[r_{t_{i+1}} \mid \mathcal{F}_{t_{i+1}}] = C + FE[r_{t_i} \mid \mathcal{F}_{t_i}], \quad Var[r_{t_{i+1}} \mid \mathcal{F}_{t_i}] = FVar[r_{t_i} \mid \mathcal{F}_{t_i}]F^T + Q,$$

which are then used as initial guesses for the state variable and its variance in the next iteration.

In this manner the estimation of the short rate is obtained. In each iteration also the prediction errors ζ_{t_i} as well as the covariance matrix of prediction errors $Var[R_{t_i} \mid \mathcal{F}_{t_{i-1}}]$ are generated. Therefore, under the assumption that the vectors of prediction errors are normally distributed, the likelihood function may be constructed as

$$\mathcal{L}(\kappa, \theta, \sigma, \lambda) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{\frac{m}{2}} \mid Var[R_{t_i} \mid \mathcal{F}_{t_{i-1}}] \mid^{\frac{1}{2}}} e^{-\frac{1}{2}\zeta_{t_i}^T Var[R_{t_i}|\mathcal{F}_{t_{i-1}}]^{-1}\zeta_{t_i}}, \tag{3.3}$$

To address the original problem of estimation of the Vasicek model parameters, the likelihood function (3.3) is then maximized according to the stated parameters κ, θ, σ and λ.

3.2. Fitting term structure of interest rates. In this subsection we briefly review the calibration approach from [8]. The procedure takes the observed yield curves R_{ij}, where i refers to days and j refers to maturities, whose values we denote by τ_j. The optimal values of the parameters, as well as the evolution of the short rate, are obtained by minimizing the objective function

$$F = \sum_{i,j} w_{ij}(R(r_i, \tau_j) - R_{ij})^2 = \sum_{i,j} \frac{w_{ij}}{\tau_j^2} \left(\log P(r_i, \tau_j) + R_{ij}\tau_j\right)^2 \tag{3.4}$$

with P and R being the bond prices and interest rates from the Vasicek model respectively, computed from (2.4)-(2.6) and w_{ij} being the weights.

However, it has to be noted that the bond prices in the Vasicek model (and therefore also interest rates which we are interested in) depend only on three independent parameters. Note that θ and λ enter the PDE (2.3) as well as the solution (2.4)-(2.6) only through the term $\kappa\theta - \lambda\sigma$. Consequently, it is possible to find a solution for the discount bond price that relies only on three parameters, i.e., σ^2 and new parameters $\alpha = \kappa\theta - \lambda\sigma$ and $\beta = -\kappa$ Parameters α, β are so-called risk-neutral parameters and they correspond to the parameters of the Vasicek model under the risk-neutral measure. It is an easy calculation to verify that using these risk-neutral parameters, the logarithm of the bond price given by (2.4)-(2.6), can be written as

$$\ln P(r, \tau) = c_0(\beta, \tau)r + c_1(\beta, \tau)\alpha + c_2(\beta, \tau)\sigma^2, \tag{3.5}$$

where functions $c_i(\beta, \tau)$ for $i = 1, 2, 3$ are given by formulae

$$c_0 = \frac{1 - e^{\beta\tau}}{\beta}, \quad c_1 = \frac{1}{\beta}\left(\frac{1 - e^{\beta\tau}}{\beta} + \tau\right), \quad c_2 = \frac{1}{2\beta^2}\left(\frac{1 - e^{\beta\tau}}{\beta} + \tau + \frac{(1 - e^{\beta\tau})^2}{2\beta}\right). \tag{3.6}$$

It follows that for a fixed β, the objective function (3.4) is a quadratic function of the parameters α, σ^2 and the values of the short rate r_i.

4. Basic concepts of solving multicriteria optimization problem. In what follows we revise the most common approaches to transforming multicriteria optimization problems. The main three approaches are taken from [13] and supplemented by their generalizations and other methods from [2]. We refer the reader to these survey papers [13], [2] and the book [12] for more details on the topic.

We consider the optimization problem in the minimization form

$$\text{minimize}_{x \in \mathbb{R}^n} F(x) = (F_1(x), F_2(x), \ldots, F_k(x))^T \tag{4.1}$$

subject to constraints

$$g_j(x) \leq 0 \text{ for } j = 1, 2, \ldots, m_1;\ h_l(x) = 0 \text{ for } l = 1, 2, \ldots, m_2. \tag{4.2}$$

Furthermore, we denote by X the set

$$X = \{x \in \mathbb{R}^n : g_j(x) \leq 0\,(j = 1, 2, \ldots, m_1); h_l(x) = 0\,(l = 1, 2, \ldots, m_2)\} \tag{4.3}$$

The most straightforward scalarization of the multicriteria problem (4.1) is the weighted sum method, when we minimize the weighted sum of the original objectives, i.e.,

$$\text{minimize}_{x \in X} \sum_{i=1}^{k} \lambda_i F_i(x), \tag{4.4}$$

where λ is a selected nonnegative vector. Recall that the set X is given by (4.3), i.e., all of the original constraints are satisfied. This can be generalized by considering a so-called partial weighting (see [17], [2]), which we will use in our calibration, when the objective functions are firstly grouped into sets with similar characteristics and each group is used to form a new objective function. Another popular method lies in keeping only one of the objectives and turning the remaining ones into constraints using a vector $\varepsilon \in \mathbb{R}^{k-1}$. If, without a loss of generality, we keep the first objective F_1, the new optimization problem reads as

$$\text{minimize}_{x \in X} F_1(x) \text{ subject to } F_i \leq \varepsilon_i\,(i = 2, \ldots, k) \tag{4.5}$$

The last approach, which we mention here in more detail, is based on creating a certain "ideal point" $y \in \mathbb{R}^k$ and trying to find a feasible solution which that $F(x)$ is as close as possible (measured by a weighted ℓ_k distance) to this ideal point. If the nonnegative weighting vector is denoted by λ, we are solving the optimization problem

$$\text{minimize}_{x \in X} \left(\sum_{i=1}^{k} \lambda_i (F_i(x) - y_i)^q\right)^{\frac{1}{q}}. \tag{4.6}$$

Note that the weighted sum approach (4.4) is a special case, when $q = 1$ a the ideal point y is zero vector. Finally, we note that there are other methods for solving multicriteria optimization problems, such as lexicographic method, goal programming methods, physical programming etc. We refer the reader to already cited references [12] , [13], [2] for a discussion on what does "minimization" of a vector in (4.1) mean, definitions of (week) Pareto optimality, proper Pareto optimality, efficiency, dominance and related concepts, their relations and necessary and sufficient conditions. Here we only note a simple property, which we will use in our application: If the vector λ in (4.4) is strictly positive and $\hat{x}$ minimizes the objective function, then there is no other feasible x such that $F_i(x) < F_i(\hat{x})$ for all indices i, i.e., there is no other feasible point which improves all the criteria considered.

5. Proposed bicriteria calibration of the Vasicek model. The main idea of our calibration is combining two criteria: minimizing squared errors in fitting term structures of interest years while giving different weights to different maturities and taking into account the probabilistic distribution of the short rate. The latter can be measured in different ways, for example evaluating the likelihood function, but we choose the Euclidean distance between the estimated short rate and the short rate obtained from the Kalman filter. The reasons are twofold. Firstly, in this way both criteria are measured in the same units and therefore they can be easily combined. Secondly, we obtain a quadratic optimization problem with a nonnegativity constraint on one of the variables, which is easy to solve.

Let us firstly denote the observed term structures of interest rates by R_{ij} where $i = 1, \dots, n$ corresponds to days and $j = 1, \dots, m$ corresponds to maturities which we denote by τ_j. Furthermore, we denote the interest rate computed from the Vasicek model computed for the short rate r and maturity τ by $R(r, \tau)$ and the corresponding bond price by $P(r, \tau)$. Following the approach from subsection 3.2 we define average (per one day) squared error of fitted interest rates with maturity τ_j as

$$f_j = \frac{1}{n} \sum_{i=1}^{n} \left(R(r_i, \tau_j) - R_{ij} \right)^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\tau_j^2} \left(\log P(r_i, \tau_j) + R_{ij} \tau_j \right)^2 . \tag{5.1}$$

Now, we define the function F_1 by weighting the functions f_j with normalized weights as

$$F_1 = \sum_{j=1}^{m} w_j f_j = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{w_j}{\tau_j^2} \left(\log P(r_i, \tau_j) + R_{ij} \tau_j \right)^2 , \text{ where } \sum_{j=1}^{m} w_j = 1 \tag{5.2}$$

As explained above, our second criterion is given by

$$F_2 = \frac{1}{n} \sum_{i=1}^{n} \left(r_i - r_i^{KF} \right)^2 , \tag{5.3}$$

where r^{KF} is the estimate of the short rate obtained by Kalman filter. Finally, combining (5.2) and (5.3) using weights $\mu \in (0, 1)$ and $1 - \mu$ respectively, we obtain the objective function

$$F = \mu F_1 + (1-\mu) F_2 = \frac{\mu}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{w_j}{\tau_j^2} \left(\log P(r_i, \tau_j) + R_{ij} \tau_j \right)^2 + \frac{1-\mu}{n} \sum_{i=1}^{n} \left(r_i - r_i^{KF} \right)^2 , \tag{5.4}$$

which is minimized with respect to parameters α, β, $\sigma^2 > 0$ and the vector of the short rate values $(r_1, \dots, r_n)^T$. In this way, we transform the multicriteria minimization of the vector $(f_1, \dots, f_m, F_2)^T$ to a two-dimensional problem $(F_1, F_2)^T$ by grouping and weighting term structure data, and finally a one-dimensional problem. Moreover, from the form of the bond price P given by (2.4)-(2.6) it follows that for a fixed β, the function (5.4) is quadratic in the remaining variables. Therefore, we firstly compute the optimal value of the objective function (5.4) for a given β and subsequently we select the optimal β (and the corresponding values of the remaining parameters).

As we only consider non-negative values for the σ^2, the problem of minimizing the objective function can be solved as a quadratic programming problem with linear constraints. Firstly, let us choose $w_j = \frac{\tau_j^2}{\sum_j \tau_j^2}$. Using the equation (3.5), the first component F_1 of F may be expanded to the form

$$\begin{aligned}
&\frac{1}{n\sum_j \tau_j^2} \sum_{i=1}^{n} \sum_{j=1}^{m} (\ln P(r_i, \tau_j) + \tau_j R_{ij})^2 = \\
&\frac{1}{n\sum_j \tau_j^2} \sum_{i=1}^{n} \sum_{j=1}^{m} \left(c_0(\beta,\tau_j) r_t + c_1(\beta,\tau_j)\alpha + c_2(\beta,\tau_j)\sigma^2 + \tau_j R_{ij}\right)^2 = \\
&\frac{1}{n\sum_j \tau_j^2} \sum_{i=1}^{n} \sum_{j=1}^{m} \left[c_0^2(\beta,\tau_j) r_i^2 + c_1^2(\beta,\tau_j)\alpha^2 + c_2^2(\beta,\tau_j)\sigma^4 + 2c_0(\beta,\tau_j)c_1(\beta,\tau_j) r_i \alpha\right. \\
&+ 2c_0(\beta,\tau_j)c_2(\beta,\tau_j) r_i \sigma^2 + 2c_1(\beta,\tau_j)c_2(\beta,\tau_j)\alpha\sigma^2 + 2c_0(\beta,\tau_j)\tau_j R_{ij} r_i \\
&\left. + 2c_1(\beta,\tau_j)\tau_j R_{ij}\alpha + 2c_2(\beta,\tau_j)\tau_j R_{ij}\sigma^2 + \tau_j^2 R_{ij}^2\right]
\end{aligned} \tag{5.5}$$

Hence, this first component of the objective function (5.4) may be rewritten to the matrix form $\frac{1}{n\sum_j \tau_j^2}\left(x^T A x + b^T x + e\right)$, where $x = (r_1, r_2, \dots, r_n, \alpha, \sigma^2)^T$ is a vector of the uknown parameters and

$$A = \begin{bmatrix}
\sum_{j=1}^m c_0^2 & 0 & \dots & 0 & \sum_{j=1}^m c_0 c_1 & \sum_{j=1}^m c_0 c_2 \\
0 & \sum_{j=1}^m c_0^2 & \dots & 0 & \sum_{j=1}^m c_0 c_1 & \sum_{j=1}^m c_0 c_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \dots & \sum_{j=1}^m c_0^2 & \sum_{j=1}^m c_0 c_1 & \sum_{j=1}^m c_0 c_2 \\
\sum_{j=1}^m c_0 c_1 & \sum_{j=1}^m c_0 c_1 & \dots & \sum_{j=1}^m c_0 c_1 & n\sum_{j=1}^m c_1^2 & n\sum_{j=1}^m c_1 c_2 \\
\sum_{j=1}^m c_0 c_2 & \sum_{j=1}^m c_0 c_2 & \dots & \sum_{j=1}^m c_0 c_2 & n\sum_{j=1}^m c_1 c_2 & n\sum_{j=1}^m c_2^2
\end{bmatrix},$$

$$b = \begin{bmatrix}
2\sum_{j=1}^m c_0 \tau_j R_{1j} \\
2\sum_{j=1}^m c_0 \tau_j R_{2j} \\
\vdots \\
2\sum_{j=1}^m c_0 \tau_j R_{nj} \\
2\sum_{i=1}^n \sum_{j=1}^m c_1 \tau_j R_{ij} \\
2\sum_{i=1}^n \sum_{j=1}^m c_2 \tau_j R_{ij}
\end{bmatrix}, \qquad e = \sum_{i=1}^{n}\sum_{j=1}^{m} \tau_j^2 R_{ij}^2.$$

In a similar way the second component F_2 of the objective function, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n}(r_i - r_i^{KF})^2 = \frac{1}{n}\sum_{i=1}^{n} r_i^2 - 2 r_i r_i^{KF} + {r_i^{KF}}^2 \tag{5.6}$$

might be expressed as $\frac{1}{n}\left(x^T C x + d^T x + f\right)$, where

$$C = diag(1,1,\dots,1,0,0), d = \left[-2r_1^{KF}, -2r_2^{KF}, \dots, -2r_n^{KF}, 0, 0\right]^T, f = \sum_{i=1}^{n}(r_i^{KF})^2.$$

In this manner the problem of finding the parameters α, σ^2 and estimation for the short rate values $(r_1,\dots,r_n)^T$ by minimization of the objective function (5.5) for a fixed value of the parameter β can be written in a following matrix form

$$\begin{aligned} &\text{minimize}_{x\in\mathbb{R}^{n+2}} \quad x^T\left[\frac{\mu}{n\sum_j \tau_j^2}A + \frac{1-\mu}{n}C\right]x + \left[\frac{\mu}{n\sum_j \tau_j^2}b + \frac{1-\mu}{n}d\right]x + g \\ &\text{subject to} \qquad \sigma^2 \geq 0. \end{aligned} \tag{5.7}$$

where $g = \frac{\mu}{n\sum_j \tau_j^2}e + \frac{1-\mu}{n}f$ is a constant, so it may be omitted when minimizing the quadratic function.

6. Application to real data. In the previous section we have proposed a calibration method, which provides the estimation of the risk-neutral parameters for the Vasicek model and also for the realizations of the short rate. To test the suggested algorithm we have used real market data. The dataset consisted of the interest rates implied by a Slovakia government bonds[1] with different lengths of maturities varying from 1 to 20 years. The dataset covered the timeframe of the second half of the year 2017.

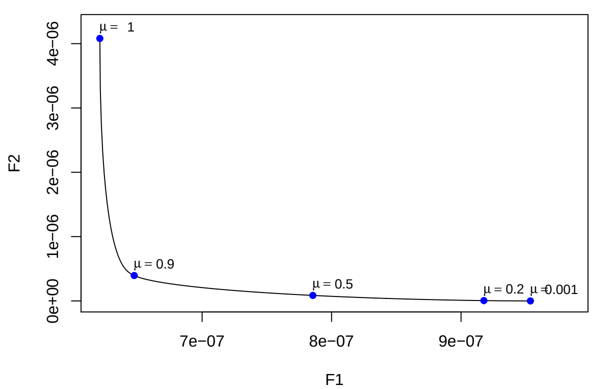

FIG. 6.1. *Graph of the relationship between $F1$ and $F2$ of the objective function (5.4), according to the different values of the parameter μ.*

We have applied the proposed algorithm on the aforementioned dataset and analyzed the behavior of the algorithm for different values of the μ. The higher the value of the μ is, the more weight is put on the first component $F1$ of the objective function. Therefore the $F1$ is a decreasing function of the μ. Analogically, the function $F2$ is increasing in the μ. Figure 6.1 shows the relationship between the values of $F1$ and

[1]source: `https://www.investing.com/rates-bonds/world-government-bonds`

$F2$, which is decreasing. One can observe a steep decline in the values of $F1$ as well as in the values of $F2$. From this point of view, it would be appropriate to choose such value of the μ that in a sense takes into consideration the minimization of both objective functions. As seen in Fig. 6.1, if μ is small (e.g. $\mu = 0.5$) then by increasing its value we obtain a significant reduction in the value of $F1$ while the value of $F2$ increases only slightly. The choice of the exact value of μ is then subjective, in our case, μ may be chosen around the value 0.9, which corresponds to the optimal values of the parameters being $\beta = 0.03951986$, $\alpha = 0.004063299$ and $\sigma^2 = 0.0001800155$.

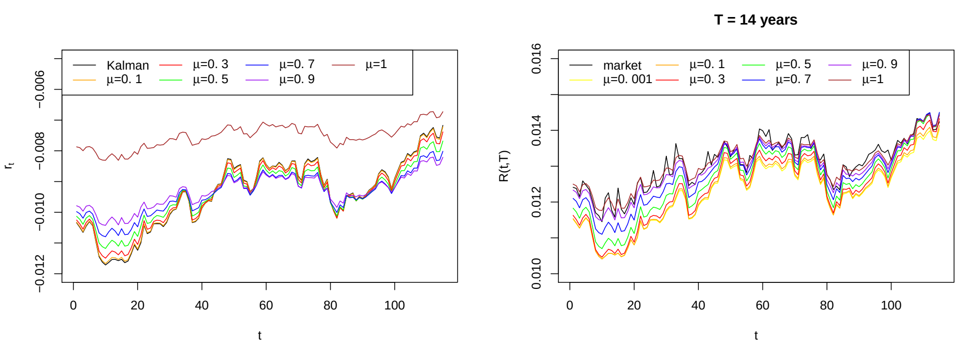

FIG. 6.2. *The left graph shows the estimated vector of the short rate gained from Kalman filtering in comparison with the estimation from the proposed algorithm for different values of μ. The right graph illustrates the fit of the estimated rate for the bond with maturity of 14 years in comparison with the real market value.*

Another way of analyzing the proposed algorithm is by looking at the estimation of the short rate. As the Kalman filtering itself generates the estimation for the latent variable, i.e., the short rate, and our algorithm also provides such estimate, we compared these two approaches, also considering different values of the μ. The left graph in Fig. 6.2 provides the graphical illustration of this comparison. It can be seen that when the $\mu = 1$, i.e., when the objective function $F2$ is not taken into the consideration, then the estimation of the short rate from the Kalman filtering differs significantly from the estimation based on the original calibration stated in section 3.2. Lowering the value of the μ just slightly results in a reduction of this difference and the estimation of the short rate starts to follow the development of the estimation from the Kalman filter. When the parameter μ is taken from the interval $(0, 0.1)$, then the estimation of the short rate from the proposed algorithm almost collides with the estimation gained from the Kalman filtering. One may also observe the fit of the estimated rate for the bond with 14 years maturity gained from the proposed algorithm in comparison with the real market values on the right graph in Fig. 6.2. According to the chosen values of the weights w_j, the fit for long term bonds is more accurate than the fit for bonds with shorter maturity. Also higher the value of μ is, the estimated rate more accurately copies the market value.

7. Conclusions. The paper deals with the calibration of the Vasicek model using multicriteria optimization. We propose a new algorithm that combines two approaches, one of which relies on fitting the term structure of the interest rates and the second one uses probability distribution of the underlying short rate. Weighted sum method is then used to scalarize the problem of minimizing two objective functions, specifically first that takes into account the weighted mean square errors between estimated and realized bond rates and second, that minimizes mean square errors between the estimations of the short rate, one of which is gained from Kalman filtering. The proposed algorithm is then tested on real market data to determine the appropriate value of the parameter μ, which for the tested dataset lies in the vicinity of 0.9. Calibration also provides an accurate fit, especially for long-term interest rates.

Possible extensions of the research include adding new criteria to the problem, choosing different method of scalarization of the multicriteria problem, or implementing the algorithm to multi-factor interest rates models, e.g., two-factor Vasicek model.

REFERENCES

[1] Y. Aït-Sahalia, *Transition densities for interest rate and other nonlinear diffusions*, The journal of finance, 54(4) (1999), pp. 1361-1395.

[2] J. S. Arora and R. Y. Marler, *Survey of multi-objective optimization methods for engineering*, Struct. Multidisc. Optim., 26 (2004), pp. 369-395.

[3] S. H. Babbs and K. B. Nowman, *Kalman Filtering of Generalized Vasicek Term Structure Models*, Journal of Financial and Quantitative Analysis, 34 (1999), pp. 115-130.

[4] D. J. Bolder, *Affine Term-Structure Models: Theory and Implementation*, Working Paper, Bank of Canada, Ottawa, 2001.

[5] D. Brigo and F. Mercurio, *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*, Springer Science & Business Media, Berlin, 2007.

[6] D. Brigo and F. Mercurio, *A deterministic shift extension of analytically–tractable and time–homogeneous short–rate models*, Finance and Stochastics, 5 (2001), pp. 369-387.

[7] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering: with MATLAB Exercises and Solutions*, John Wiley & Sons Inc., Hoboken, 1996.

[8] Z. Bučková, J. Halgašová, and B. Stehlíková, *Estimating the short rate from the term structure in the Vasicek model*, Tatra Mountains Mathematical Publications, 61 (2014), pp. 87-103.

[9] Z. Bučková, J. Halgašová, and B. Stehlíková, *Short Rate as Sum of Two CKLS-Type Processes*, Numerical Analysis and Its Applications, (2017), pp. 243-251.

[10] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, *An empirical comparison of alternative models of the short-term interest rate*, The journal of finance, 47(3) (1992), pp. 1209-1227.

[11] T. Corzo Santamaria and E. S. Schwartz, *Convergence within the EU: Evidence from interest rates*, Economic Notes, 29(2) (2000), pp. 243-266.

[12] M. Ehrgott, *Multicriteria optimization (Vol. 491)*, Springer Science & Business Media, 2005.

[13] M. Ehrgott, *Multiobjective optimization*, Ai Magazine, 29(4) (2008), pp. 47-57.

[14] T. Fischer, A. May, and B. Walther, *Fitting Yield Curve Models Using the Kalman Filter*, Proceedings in Applied Mathematics and Mechanics, 3 (2003), pp. 507-508.

[15] J. James and N. Webber, *Interest Rate Modelling*, John Wiley & Sons Inc., Hoboken, 2000.

[16] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transaction of the ASME—Journal of Basic Engineering, 82(Series D) 1960, pp. 35-45.

[17] J. Koski and R. Silvennoinen, *Norm methods and partial weighting in multicriterion optimization of structures*, International Journal for Numerical Methods in Engineering, 24(6) (1987), pp. 1101-1121.

[18] P. S. Maybeck, *Stochastic models, estimation, and control, Volume 1*, Academic Press, New York, 1979.

[19] K. Mikula, B. Stehlíková, and D. Ševčovič, *Analytical and numerical methods for pricing financial derivatives*, Nova Science, Hauppauge, 2011.

[20] K. B. Nowman, *Continuous-time short term interest rate models*, Applied Financial Economics, 8(4) (1998), pp. 401-407.

[21] H. W. Sorenson, *Least-Squares estimation: from Gauss to Kalman*, IEEE Spectrum, 7 (1970), pp. 63-68.

[22] D. Ševčovič and A. Urbánová - Csajková, *On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model*, Central European Journal of Operations Research, 13(2) (2005).

[23] O. Vašíček, *An equilibrium characterization of the term structure*, Journal of financial economics, 6 (1977), pp. 177-188.