KYBERNETIKA — VOLUME 48 (2012), NUMBER 3, PAGES 550–566

COPULA APPROACH TO RESIDUALS OF REGIME-SWITCHING MODELS

Anna Petričková and Magda Komorníková

The autocorrelation function describing the linear dependence is not suitable for description of residual dependence of the regime-switching models. In this contribution, inspired by Rakonczai ([20]), we will model the residual dependence of the regime-switching models (SETAR, LSTAR and ESTAR) with the autocopulas (Archimedean, EV and their convex combinations) and construct improved quality models for the original real time series.

Keywords: autocopula, time series, residuals, regime-switching models

Classification: 93E12, 62A10

1. INTRODUCTION

The first models used for modelling economical and financial time series had a linear character (shocks were assumed to be non-correlated but not necessarily independent and identically distributed – iid). Although many of the models commonly used in empirical finance are linear, the nature of financial data suggests that nonlinear models are more appropriate [8].

Therefore, in recent years there were proposed many time series models, which formalize the idea of the different regimes existence in time series. These models can be used for the modelling of financial yields, hydrological and geodetic time series, and so on ([1, 4, 13]). They have the nonlinear character.

In this work we focus on the models from the class of regime-switching models where changes in variability are related to (or predicted by) recent past values of the observed time series. We restrict our attention to the models where the dynamic behavior of the time series can be described adequately by a linear AR model in each of the regimes. We deal with three classes of the regime-switching models, regimes of which are determined by observable variables (SETAR, LSTAR and ESTAR).

The paper is organized as follows. After a general introduction, the theoretical basis of the regime-switching models with regimes determined by observable variables, copulas and some tests are described. The paper continues with the application of the residual dependence modelling for the regime-switching models (SETAR, LSTAR and ESTAR) with autocopulas and the constructing of improved quality models for the original real time series.

2. THEORETICAL BASIS

2.1. Overview of regime-switching models with regimes determined by observable variables

Typical models belonging to this class are TAR models ("**T**hreshold **A**uto**R**egressive"). They are well to interpret and also very suitable for modelling a lot of real data. They form the basis of the regime-switching models with regimes determined by observable variables.

These models assume that any regime in time t can be given by any observed variable q_t (indicator variable). Values of q_t are compared with threshold value c.

2.1.1. SETAR model

The special case arises when q_t is taken to be a lagged value of the time series itself that is, $q_t = X_{t-d}$ for a certain integer $d > 0$. The resulting model is called a **S**elf-**E**xciting **T**hreshold **A**uto**R**egressive (SETAR) model. For example the 2-regime SETAR model with AR(p) in both regimes has form

$$\begin{aligned} X_t &= (\phi_{0,1} + \phi_{1,1}X_{t-1} + \cdots + \phi_{p,1}X_{t-p})\,[1 - \mathbf{1}(X_{t-d} > c)] \\ &+ (\phi_{0,2} + \phi_{1,2}X_{t-1} + \cdots + \phi_{p,2}X_{t-p})\,\mathbf{1}(X_{t-d} > c) + \varepsilon_t \end{aligned} \tag{1}$$

where $\{\varepsilon_t\}$ is the strict white noise process with $E[\varepsilon_t] = 0$, $D[\varepsilon_t] = \sigma_\varepsilon^2$ for all $t = 1, \ldots, T$ and $\mathbf{1}$(A) is the *indicatorfunction* with values $\mathbf{1}$(A) = 1 if the event A occurs and $\mathbf{1}$(A) = 0 otherwise.

In the case of a 3-regime model we have to define four constants $c_i,\ i = 0, 1, 2, 3,$ where $-\infty = c_0 < c_1 < c_2 < c_3 = \infty$. The SETAR model with AR(p) in all regimes has the form

$$X_t = \phi_{0,j} + \phi_{1,j}X_{t-1} + \cdots + \phi_{p,j}X_{t-p} + \varepsilon_t \tag{2}$$

if $c_{j-1} < X_{t-d} \leq c_j$, j = 1, 2, 3.

For more details see Arlt and Arltová ([2]), Frances and Van Dijk ([8]).

2.1.2. STAR model

If we replace the indicator function $\mathbf{1}(q_t > c)$ by a continuous function $F(q_t, \delta, c)$ (so called the *transition function*, where δ is the smooth parameter) which changes smoothly from 0 to 1 as q_t increases, the resulting model is called **S**mooth **T**ransition **A**uto**R**egressive (STAR) model ([2, 8]). For example the 2-regime STAR model with AR(p) in both regimes and the indicator variable $q_t = X_{t-d}$ has the form

$$\begin{aligned} X_t &= (\phi_{0,1} + \phi_{1,1}X_{t-1} + \cdots + \phi_{p,1}X_{t-p})\,[1 - F(X_{t-d}, \delta, c)] \\ &+ (\phi_{0,2} + \phi_{1,2}X_{t-1} + \cdots + \phi_{p,2}X_{t-p})\,F(X_{t-d}, \delta, c) + \varepsilon_t \end{aligned} \tag{3}$$

If the transition function $F(q_t, \delta, c)$ is *the logistic function*

$$F(q_t, \delta, c) = \left(\frac{1}{1 + \exp\left(-\delta\,(q_t - c)\right)}\right), \quad \delta > 0, \tag{4}$$

the resulting model is called a **L**ogistic **STAR** (LSTAR) model.

If the transition function $F(q_t, \delta, c)$ is *the exponential function*

$$F(q_t, \delta, c) = 1 - \exp\left(-\delta\,(q_t - c)^2\right), \quad \delta > 0, \tag{5}$$

the resulting model is called an **E**xponential **STAR** (ESTAR) model. For more details on STAR models see, e. g. Arlt and Arltová [2], Frances and Van Dijk [8], Teräsvirta [21]).

2.2. Empirical specification procedure for nonlinear models

In the process of the empirical specification for the nonlinear models the following steps are recommended [11]:

1. specify an appropriate linear AR model of order p for the time series under investigation,
2. test the null hypothesis of linearity against the alternative of SETAR- and/or STAR-type nonlinearity; this step also consists of selecting the appropriate variable that determines the regimes,
3. estimate the parameters in the selected model,
4. evaluate the model using diagnostic tests,
5. modify the model if necessary,
6. use the model for descriptive or forecasting purposes.

When we test the null hypothesis of linearity against the alternative of SETAR-type nonlinearity, we need to know estimates of parameters for the nonlinear model. Similarly, we test optimal 2–regime models (with the minimum value of BIC criterion) for remaining nonlinearity against three-regime alternatives. (For details of testing procedures see e. g. [8]). The residuals of the selected models are further tested for serial correlations (see [8]). The residuals of the regime-switching models are supposed to be independent (not only serially non-correlated). This property can be tested e. g. by the BDS test [3].

2.3. The BDS test

There are number of reasons why the BDS statistic has become so widely used. It can be applied as a goodness of fit test to any model that can be transformed into model with additive i.i.d. errors and whose parameters can be estimated $\sqrt{T}$-consistently (where T is the length of time series). Also, this statistic can be used to test for stochastic linearity. The other reason is, that the asymptotic distribution theory of the BDS statistic does not require higher moments to exist. This is important in financial economics because of the problems that heavy–tailed distributions can cause when using most of the standard test statistics. This test was presented in the of Brock, Dechert, Scheinkman, and Le Baron [3] and can be used to test the independence in residuals $\{\hat{e}_t\}$.

For some $n \in N$ and $\varepsilon > 0$ the test is based on the correlation integral

$$C_{n,\varepsilon} = \frac{\sum\sum_{m+1 \leqq \tau \leqq T_n} \mathbf{1}\left(\|\widehat{\mathbf{e}}_{\mathbf{t},\mathbf{n}} - \widehat{\mathbf{e}}_{\tau,\mathbf{n}}\| < \varepsilon\right)}{2[(T_n - 1)T_n]} \tag{6}$$

where $T_n = T - n + 1, \hat{\mathbf{e}}_{\mathbf{t},\mathbf{n}} = (\hat{e}_t, \ldots, \hat{e}_{t+n-1})$, $\mathbf{1}(\text{A})$ is the indicator of the event A, and $\|\cdot\|$ denotes the maximum norm (also known as Chebyshev norm) in $\Re^d$ (i. e., $\|z\| = \max_{1<i<d} |z_i|$ for $z = (z_1, \ldots, z_d)'$). Then the BDS statistic is

$$\Lambda_{BDS} = \frac{\sqrt{(T-m)}(C_{n,\varepsilon} - C^n_{1,\varepsilon})}{\sigma_{n,\varepsilon}} \tag{7}$$

where

$$\frac{1}{4}V_{n,\varepsilon} = K^n_\varepsilon + (n-1)^2 C^{2n}_\varepsilon - n^2 K_\varepsilon C^{2(n-1)}_\varepsilon + 2\sum_{j=1}^{n-1} K^{n-1}_\varepsilon C^{2j}_\varepsilon,$$

$$K_\varepsilon = \frac{1}{(T-m)^3} \sum_{K=m+1}^{T} \sum_{\tau=m+1}^{T} \sum_{t=m+1}^{T} \mathbf{1}\left(|\hat{e}_\kappa - \hat{e}_\tau| < \varepsilon\right) \mathbf{1}\left(|\hat{e}_\tau - \hat{e}_t| < \varepsilon\right),$$

$$C_\varepsilon = \frac{1}{(T-m)^2} \sum_{\tau=m+1}^{T} \sum_{t=m+1}^{T} \mathbf{1}\left(|\hat{e}_\tau - \hat{e}_t| < \varepsilon\right),$$

T is length of time series and m is order of process AR. The test is implemented here using "embedding dimension" $n = 2$ (number of time-lagged residuals) and "metric bound" $\varepsilon = \tilde{\sigma}$, where $\tilde{\sigma}^2 = \frac{1}{T-m-1}\sum_{t=m+1}^{T} \hat{e}^2_t$; the latter choice is in the centre of the region $\frac{1}{2}\tilde{\sigma} \leq \varepsilon \leq \frac{3}{2}\tilde{\sigma}$, that is recommended in Brock et al. [3].

Λ_{BDS} has N(0,1) asymptotic distribution when $\{e_t\}$ are i.i.d.

When BDS test at the significant level shows residual dependence, we use n-lag autocopulas for modelling of these dependent residuals.

Inspired by the approach of Rakonczai [20] we have applied autocopulas to the time series of the residuals exhibiting nonlinear dependences in order to gauge how much they violate the assumptions of independence. We have tried to describe the dependences of the (time lag 1) residuals of regime-switching models by means of copulas.

2.4. Copula

A 2-dimensional copula is the function ([14, 17]) $C : [0,1]^2 \longrightarrow [0,1]$, which satisfies

- the boundary conditions:

  $C(0,y) = C(x,0) = 0$ and $C(1,y) = y, C(x,1) = x$ for all $x, y \in [0,1]$,

- the 2-increasing property:

  $C(x_1,y_1) + C(x_2,y_2) - C(x_1,y_2) - C(x_2,y_1) \geq 0$

  for all $x_1, y_1, x_2, y_2 \in [0,1]$ such that $x_1 \leq x_2$ and $y_1 \leq y_2$.

The most important applications of the 2-dimensional copulas are related to a well-known and very convenient rewriting of the joint distribution function F of a 2-dimensional random vector (X, Y) in the form

$$F(x,y) = C(F_X(x), F_Y(y)) \tag{8}$$

where F_X, F_Y are marginal distribution functions.

Rakonczai et al. [20] introduced autocopulas to describe the lag self-dependence structure of a time series. Given a strictly stationary time series X_t and set of lags $\mathcal{L} = \{d_i \in Z^+ : i = 1, \dots, k\}$, the *autocopula* $C_{X,\mathcal{L}}$ is defined as the copula on the $k+1$ dimensional random vector $(X_t, X_{t-d_1}, \dots, X_{t-d_k})$. The *k-lag autokopula* $C_{X,k}$ is the autocopula with lag $k \in Z^+$. In other words, autocopula is an ordinary copula related to original and the lagged time series, and as such it describes the interdependence structure in more detail than autocorrelation does, specifically, it takes into account non-linear interdependencies as well.

2.4.1. Some bivariate copulas

Archimedean class of copula

Copula C belongs to the Archimedean class if (see e. g. [6, 14, 17])

$$C_\phi(x,y) = \phi^{(-1)}(\phi(x) + \phi(y)) \quad \text{for } x, y \in (0,1],$$

where $\phi : (0,1] \to [0,\infty)$ is a convex, decreasing function (satisfying $\phi(1) = 0$) that is called a *generator* of the copula C_ϕ, and $\phi^{(-1)} : [0,\infty) \to [0,1]$ is given by

$$\phi^{(-1)}(x) = \sup\{t \in (0,1] \mid \phi(t) \geq x\} = \begin{cases} \phi^{-1}(x) & x < \phi(0^+) \\ 0 & \text{else.} \end{cases}$$

Let $(X,Y)'$ be a vector of continuous random variables with copula C. Then Kendall's tau for $(X,Y)'$ is given by

$$\tau(X,Y) = 4 \int\int_{[0,1]^2} C(u,v)\,\mathrm{d}C(u,v) - 1.$$

As a generator uniquely determines an Archimedean copula, different choices of generators yield many families of copulas that consequently, besides the form of the generator, differ in the number and the range of parameters. We summarize some basic facts related to the most important 1- and 2-parameter families of Archimedean copulas ([6, 17]). Note, that Clayton and Gumbel copulas model only positive dependence (measured by Kendall's τ), while Frank copulas cover the whole range [-1,1].

Some examples:

- Gumbel family

  $C_\theta(x,y) = \exp^{-\left((-\ln x)^\theta + (-\ln y)^\theta\right)^{\frac{1}{\theta}}}$, where $\theta \geq 1$,
  $\tau = \frac{\theta - 1}{\theta}$.

- Strict Clayton family (Kimeldorf and Sampson)

  $C_\theta(x,y) = \left(x^{-\theta} + y^{-\theta} - 1\right)^{\frac{-1}{\theta}}$ and $C_0(x,y) = \Pi\,(x,y) = xy$,

  $\tau = \frac{\theta}{\theta+2}$ where $\theta > 0$.

- Frank family

  $C_\theta(x,y) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta x}-1)(e^{-\theta y}-1)}{(e^{-\theta}-1)}\right)$, where $\theta \in \Re$,

  $\tau = 1 - \frac{4}{\theta}\left(1 - D_1(\theta)\right)$,

  $D_1(z) = \frac{1}{z}\int_0^z \frac{t}{e^t-1}\,\mathrm{d}t$ is the Debye function.

- Joe's BB1 family

  $C_{\theta,a}(x,y) = \left(1 + \left((x^{-a}-1)^{\theta} + (y^{-a}-1)^{\theta}\right)^{\frac{1}{\theta}}\right)^{\frac{-1}{a}}$,

  $\tau = 1 - \frac{2}{\theta(a+2)}$ where $\theta \geq 1$ and $a > 0$.

A rich overview of Archimedean copulas is presented in Embrechts et. al [6], Genest and Favre [9], Joe [14] and Nelsen [17].

Extreme Value (EV) Copulas

Their characteristic relation is

$$C\left(x^t, y^t\right) = \left(C(x,y)\right)^t$$

for any $t > 0$ and $(x,y) \in [0,1]^2$.

A bivariate copula C is an extreme-value copula if and only if it can be represented in the form

$$C(x,y) = \exp\left(\ln(xy) A\left(\frac{\ln(x)}{\ln(xy)}\right)\right)$$

where the function $A : [0,1] \to [1/2, 1]$ is a convex function and satisfies the inequality

$$\max(t, 1-t) \leq A(t) \leq 1$$

for all $t \in [0,1]$. The function A is called a *dependencefunction*.

We have worked with the following types of EV copulas [10]:

- Gumbel A family (Mixed model)

  $C_\theta(x,y) = xy\exp\left(-\theta\frac{\ln(x)\ln(y)}{\ln(x)+\ln(y)}\right)$,

  $\tau = \frac{8\arctan\sqrt{\frac{\theta}{4-\theta}}}{\sqrt{\theta(4-\theta)}} - 2$, where $0 \leq \theta \leq 1$.

- Galambos family

  $C_\theta(x,y) = xy\exp\left(\left((-\ln x)^{-\theta} + (-\ln y)^{-\theta}\right)^{-\frac{1}{\theta}}\right)$,

  $\tau = \frac{\theta+1}{\theta}\int_0^1 \left(\frac{1}{s^{1/\theta}} + \frac{1}{(1-s)^{1/\theta}} - 1\right)^{-1} \mathrm{d}s$, where $\theta \geq 0$.

Notice that the only class of associative EV copulas are just Gumbel copulas (with $Min(x,y) = \min\{x,y\}$ as it's limiting case when the parameter of copula from Gumbel family approaches infinity).

Convex combinations of Copulas

A very useful class for fitting the investigated couples of time series has been obtained in the class of convex combinations of copulas C_1 and C_2 given by

$$C_\alpha(x,y) = \alpha C_1(x,y) + (1-\alpha)C_2(x,y)$$

for $\alpha \in [0,1]$.

In practical data fitting by the copulas of this form we have applied the MLE method (with initial parameters estimate received by the minimalization of the mean square distances to the fitted empirical copulas). In order to reduce complexity of required computations we have used convex combinations of Archimedean copulas only.

2.4.2. Goodness of fit test for copulas

Let $\{(x_j,y_j), j = 1,\dots,T\}$ be T modeled 2-dimensional observations, F_X, F_Y their marginal distribution functions and F their joint distribution function.

We say that the class of copulas C_θ is correctly specified if there exists θ_0 so that holds

$$F(x,y) = C_{\theta_0}\left(F_X(x), F_Y(y)\right).$$

Let c_θ be the density function of C_θ, copula C_θ has to be absolutely continuous, ∇_θ be the vector of all first-order partial derivatives of $\ln c_\theta\left(F_X(x), F_Y(y)\right)$ and ∇^2_θ be the square matrix of second-order partial derivatives of the same function $\ln c_\theta\left(F_X(x), F_Y(y)\right)$ (the Hessian matrix or simply the Hessian) are continuous on the domain of their parameters. White [23] showed that under correct specification of the copula class C_θ the following information matrix equivalence holds:

$$-\mathbf{A}_{\theta_0} = \mathbf{B}_{\theta_\mathbf{0}},$$

where

$$\mathbf{A}_\theta = E[\nabla^2_\theta \ln c_\theta\left(F_X(X), F_Y(Y)\right)]$$

$$\mathbf{B}_\theta = E[\nabla_\theta \ln c_\theta\left(F_X(X), F_Y(Y)\right) \nabla'_\theta \ln c_\theta\left(F_X(X), F_Y(Y)\right)].$$

The proposed procedure [18] is based on the empirical distribution functions

$$\hat{F}_X(s) = \frac{1}{T}\sum_{i=1}^{T} \mathbf{1}(x_i \le s), \hat{F}_Y(s) = \frac{1}{T}\sum_{i=1}^{T} \mathbf{1}(y_i \le s)$$

and also on a consistent estimator $\hat{\theta}$ of θ_0 that maximizes $\ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right)]$. To introduce the sample versions of $\mathbf{A}$ and $\mathbf{B}$ put

$$\mathbf{A}_i(\theta) = \nabla^2_\theta \ln c_\theta\left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right)$$

$$\mathbf{B}_i(\theta) = \nabla_\theta \ln c_\theta \left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right) \nabla'_\theta \ln c_\theta \left(\hat{F}_X(x_i), \hat{F}_Y(y_i)\right)$$

$$\hat{\mathbf{A}}_\theta = \frac{1}{T}\sum_{i=1}^{T} \mathbf{A}_i(\theta)$$

$$\hat{\mathbf{B}}_\theta = \frac{1}{T}\sum_{i=1}^{T} \mathbf{B}_i(\theta)$$

and

$$\mathbf{d}_i(\theta) = vech\left(\mathbf{A}_i(\theta) + \mathbf{B}_i(\theta)\right)$$

where vech($\mathbf{M}$) is the vector of dimension [k(k+1)/2] containing the upper triangle (in the lexicographic ordering) of the symmetric matrix $\mathbf{M}$ of the type $k \times k$ (where k is the space dimension of parameters θ).

Put $\hat{\mathbf{D}}_\theta = \frac{1}{T}\sum_{i=1}^{T} \mathbf{d}_i(\theta)$. Under the hypothesis of proper specification the statistics $\sqrt{T}\ \hat{\mathbf{D}}_\theta$ has asymptotical distribution $N(0,\mathbf{V})$, where $\mathbf{V}$ is estimated by $\hat{\mathbf{V}} = \frac{1}{T-1}\sum \mathbf{d'}_i(\theta)\mathbf{d}_i(\theta)$.

Therefore

$$\xi = T\hat{\mathbf{D}}'_\theta \hat{\mathbf{V}}_\theta^{-1}\hat{\mathbf{D}}_\theta \tag{9}$$

asymptotically follows $\chi^2_{\frac{k(k+1)}{2}}$.

2.4.3. Fitting of copulas

In practical fitting of the data we have utilized the *maximum pseudolikelihood method* (MPLE) of parameter estimation (with initial parameters estimates received by the minimalization of the mean square distance to the empirical copula C_n presented e. g. in Genest and Favre [9]). It requires that the copula $C_\theta(x,y)$ is absolutely continuous with density $c_\theta(x,y) = \frac{\partial^2}{\partial x\, \partial y} C_\theta(x,y)$. This method (described e. g. in [9]) involves maximizing a rank-based log-likelihood of the form

$$L(\theta) = \sum_{i=1}^{T} \ln\left(c_\theta\left(\frac{R_i}{T+1}; \frac{S_i}{T+1}\right)\right)$$

where T is the sample size, R_i stands for the rank of X_i among $X_1,\dots,X_T$, S_i stands for the rank of Y_i among $Y_1,\dots,Y_T$ and θ is vector of parameters in the model. Note that arguments $\frac{R_i}{T+1}$, $\frac{S_i}{T+1}$ equal to the corresponding values of the empirical marginal distributional functions of random variables X and Y.

To compare a goodness of fit of our estimated model we have used *Takeuchi criterion TIC* ([12]) that is a robustified version of the Akaike criterion *AIC*

$$TIC = -2L(\theta) + 2Tr\left(\hat{\mathbf{B}}_\theta \cdot \hat{\mathbf{A}}_\theta^{-1}\right). \tag{10}$$

Smaller *TIC* value means an improvement of the quality of model fitting.

2.5. Forecast

Our main goal is to evaluate a regime-switching model by comparing of out-of-sample forecast to observed values. Following [8] and [11], let g denote the model describing X_t through formula

$$X_t = g(\Omega_{t-1}; \theta) + e_t \ , \tag{11}$$

with history Ω_{t-1} of the time series up to and including X_{t-1}, parameter vector θ and i.i.d. white noise e_t at time t with distribution D . For the sake of demonstration assume a general nonlinear autoregressive model of order one. The optimal one-step-ahead forecast is

$$\hat{X}_{t+1|t} = \mathrm{E}[X_{t+1}|\Omega_t] = g(X_t; \theta) \tag{12}$$

which can be achieved with no difficulty. The two-step case is not easy,

$$\hat{X}_{t+2|t} = \mathrm{E}[X_{t+2}|\Omega_t] = \mathrm{E}[g(X_{t+1}; \theta)|\Omega_t], \tag{13}$$

because the linear conditional expectation operator E cannot be interchanged with the nonlinear operator g. It helps to link both forecasts (12) and (13),

$$\hat{X}_{t+2|t} = \mathrm{E}[g(g(X_t; \theta) + e_{t+1}; \theta)|\Omega_t] = \mathrm{E}[g(\hat{X}_{t+1|t} + e_{t+1}; \theta)|\Omega_t]. \tag{14}$$

If we do not want to ignore the random term e_{t+1} (that would be a naïve forecast) we might attempt to obtain the conditional expectation directly by computing

$$\hat{X}_{t+2} = \int_{-\infty}^{\infty} g(\hat{X}_{t+1|t} + \epsilon)\, \mathrm{d}\Phi(\epsilon), \tag{15}$$

where $\Phi(\epsilon)$ is the distribution function of D. Instead of numerical integration we approximate it by Monte-Carlo method,

$$\hat{X}_{t+2|t} = \frac{1}{N} \sum_{i=1}^{N} g(\hat{X}_{t+1|t} + \epsilon_i; \theta), \tag{16}$$

where N is some large number and ϵ_i, $i = 1, \ldots, N$, are random numbers drawn from distribution D. In practise the function g is not known, it has to be specified and estimated, so that e_t is not usually a white noise and has temporal relationships. These are most probably of nonlinear quality, so it seems natural to employ copulas to simulate ϵ_i. Given residual at time t, $\hat{e}_t = X_t - \hat{X}_{t|t-1}$ with associated distribution function Φ, ϵ for next-step-ahead forecast can be obtained as quantile of conditional distribution function

$$F_{Y|X=x}(y) = \frac{\partial C(u, F_Y(y))}{\partial u}\Big|_{u=F_X(x)}, \tag{17}$$

that is

$$\epsilon_i = F^{-1}_{Y|X=e_t}(p), \tag{18}$$

where p is a random number drawn from uniform distribution $U(0,1)$. Recall that C is a bivariate copula and F_X, F_Y are marginal distribution functions, in our case assumed to identicaly belong to zero-mean normal distribution $N(0, \hat{\sigma}^2)$ with variance estimated from $\hat{e}_t$. Because of successively applied one-step prediction, 1-lag autocopula is chosen as the copula C. Since the usual inference theory for copulas assumes i.i.d. observations, as pointed out in [19], one should not use all single pairs of observations $\{(X_{i-1}, X_i) : i \in \{2, \ldots, T\}\}$ for copula fitting (which in most cases are not independent), instead a thined subset is recommended. On this account we took every mth pair (with m presumably large enough).

3. REVIEW OF RESULTS

This section was inspired by Komorník and Komorníková [15]. Below we summarize all results in tables and graphs. For the investigations there have been used 50 real data series (exchange rates, various macroeconomic data and other financial data series) and for all calculations the system *MATHEMATICA*, version 7 has been used. For all statistical tests used in the subsequent analyses, we have considered the significance level 0.05. We have omitted 5 the most recent values from each of the considered time series (they were used for checking of the predictive properties of the resulting models).

For the considered time series and models we have sequentially performed the following succession of procedures:

1. At first, we 'fitted' these real time series with SETAR, LSTAR and ESTAR model [8]. In each class we have selected the best model (optimizing the number of states and the order of the local autoregressive models) on the basis of the BIC criterion ([8]).

2. For every time series we have chosen the best model from model classes mentioned above; as a criterion we have used the lowest values of the residual standard deviation.

3. We have applied the serial autocorrelation test of the residuals for the selected models.

4. The residuals of these models have been supposed to be not only serially non-correlated but also independent (which can be tested by the BDS test [3]).

5. We have applied autocopulas to the time series of the above mentioned residuals. We have tried to describe the dependencies of the (time lag 1) residuals of selected models by means of copulas (EV, Archimedean and their convex combinations). For each couple $(\hat{e}_t, \hat{e}_{t-1})$, $t \in \{2, 2+m, \ldots, 2+r\,m \leq T\}$, where m is the thinning parameter (presumably large enough) (see [20]) and each class of copulas we have sequentially performed the following procedures:

   - calculation of Maximum pseudo–likelihood estimates and TIC,
   - calculation of p-values of the goodness of fit tests for all candidate copula models and selection only of the models that pass GoF tests,
   - selection of the optimal models with the minimal TIC criterion (from models that pass GoF tests).

3.1. Selection of the best model

At first, we have tested linearity against the (2-regime) SETAR, LSTAR, or ESTAR type of nonlinearity. For 5 illustrative time series

- TS1 = Final consumption households, Euro area 16, seasonally adjusted (1995Q1 – 2009Q4),
- TS2 = Industrial production index USA, monthly seasonally adjusted data (01/1997 – 12/2009)
- TS3 = UK march 2010, exchange rates GBP/EUR, monthly data (03/1999 – 03/2010)
- TS4 = Danish krone, exchange rates DKK/EUR, monthly data (03/1999 – 03/2010)
- TS5 = Unemployement USA, monthly seasonally adjusted data (06/1999 – 06/2010)

the corresponding results are included in the Table 1. The first time series in the table are modeled with SETAR, the second and third one with LSTAR and the last two with ESTAR model.

Time series	**The linearity test**
TS1	0.007406
TS2	0.000106
TS3	0.000004
TS4	$< 10^{-6}$
TS5	0.000480

Tab. 1. The p-values for the test of linearity of estimated models.

Next, we have continued by testing remaining nonlinearity of optimal 2-regime SETAR, LSTAR, or ESTAR models (with the minimum value of the *BIC* criterion) against alternatives of 3-regime model. The corresponding results are included in the second column of Table 2.

Time series	**The remaining nonlinearity test**	**The serial correlation test**	**The BDS test**
TS1	0.0138	**0.3631**	0.023
TS2	0.00671	**0.9766**	$< 10^{-6}$
TS3	0.00001	**0.7074**	0.00005
TS4	$< 10^{-6}$	**0.8175**	0.00084
TS5	**0.1373**	**0.5411**	0.0102

Tab. 2. The p-values for the diagnostic checking of estimated models.

We have chosen four 3-regimes models, only the last illustrative time series is described by 2-regimes model.

Next we have applied the residual autocorrelation tests for the selected model's outcomes (see the third column of Table 2 for the results). However, the residuals of the models are supposed to be not only serially non-correlated but also independent. This property has been tested by the BDS test of independence (the fourth column of Table 2).

In the Table 3, there is a summary of the test results – the number of the time series with dependent and non-correlated residuals (for all models).

Type of model	**Number**
SETAR	9
LSTAR	16
ESTAR	20

Tab. 3. Summary of the tests results – dependent & non-correlated.

3.2. The best copulas

For each couple $(\hat{e}_t, \hat{e}_{t-1})$ and each class of copulas we have subsequently performed the sequence of procedures that we have described above and we have selected the optimal models that attain the minimum value of the TIC criterion.

The results for 5 illustrative time series are presented in Table 4. We see that none of the selected optimal models have been rejected by the GoF test. Moreover, significant results of the BDS test justify efforts to model dependencies between the residuals (despite low values of the Kendall τ's).

Time series	Kendall's τ	Type of copulas	p–values	α	θ_1	θ_2	a
TS1	-0.03774	Frank& Joe	0.10897	0.90041	-0.74241	1	2.85563
TS2	-0.07719	Clayton & Frank	0.23334	0.38532	0.66171	-2.11948	x
TS3	-0.00698	Gumbel A	0.16552	x	0.11818	x	x
TS4	0.08917	Frank	0.31532	x	0.82413	x	x
TS5	-0.00574	Gumbel & Joe	0.23882	0.82887	1.05539	13.4784	4.79458

Tab. 4. Optimal copulas for the pairs of residuals.

3.3. Modified models

In this section we aim to improve classical SETAR, LSTAR and ESTAR models. Instead of $\{e_t\}$ (which is the strict white noise process with $E[e_t] = 0, D[e_t] = \sigma_e^2$ for all $t = 1, \ldots, T$), we have used the autocopulas that we have chosen as the best copulas above (for each real time series). To compare the quality of the optimal models without (first category) and with (second category) application of copulas for modelling of their residuals, we have compared their prediction mean square error *MSE* for the differences between the values of the (1-step-ahead) forecasts of the optimal models and the last 5 values of empirical data as well as standard deviations (σ_e). In the case of models in the second category, we have applied the fitted copulas for the residuals of the optimal models in the calculations of the forecasts.

In the columns in Table 5 we can see the standard deviations σ_e of the residuals in both categories without and with copulas and also their percentage changes.

Time series	σ_e for model without copulas	σ_e for model with copulas	Relative reduction (%)
TS1	0.048415	0.031102	35.76
TS2	0.000517	0.000267	48.44
TS3	$4.19 * 10^{-6}$	$1.97 * 10^{-6}$	53.04
TS4	0.000032	0.000015	53.21
TS5	0.036251	0.020662	43.01

Tab. 5. The standard deviation of residuals for the model without and with copulas.

In the next Table 6 there are for the same 5 illustrative time series prediction errors *MSE* and *MAE* (Mean Absolute Error) for 1-step-ahead forecasts for 5 time units of residuals in both categories without and with copulas and also their percentage changes.

Time series	MSE without c.	MSE with c.	Relative reduction (%)	MAE without c.	MAE with c.	Relative reduction (%)
TS1	1.33668	1.40617	-5.20	0.93928	0.97576	-3.88
TS2	0.01584	0.00850	46.33	0.12036	0.06507	45.33
TS3	$1.79 * 10^{-6}$	$2.11 * 10^{-6}$	-16.98	0.00117	0.00132	-12.63
TS4	0.000018	$5.22 * 10^{-6}$	71.03	0.00363	0.00189	47.94
TS5	0.05432	0.06554	-20.67	0.20131	0.21666	-7.63

Tab. 6. The prediction errors for 1-step-ahead forecasts for 5 time units of residuals for model without and with copulas.

In order to compare visually the improvements in standard deviations of the residuals for the 2 categories of models, we present graphs comparing the original data (marked by dots) of 5 illustrative time series by the optimal models from the first (dashed line) and the second category (solid line) – see the Figure 1.

In order to compare visually forecasting performance of the 2 categories of models, we show (in Figure 2) the differences in fitting of the last 5 values of the illustrative time series (marked by *) by 1-step-ahead forecasts by the optimal models from the first (dashed line) and the second category (solid line).

The average improvement in description for 31 modeled time series is 51.42 % but in the case of prediction it comes to the small degradation in almost all cases. Average degradation in prediction is -11.81 %. It is necessary to remark that most nonlinear techniques give good in-sample fits to real time series, but they are usually outperformed by models of random walks or random walks with drift when are used for forecasting [5]. Next we directly quote Enders and Pascalau [7]: "Teräsvirta [22] summarizes much

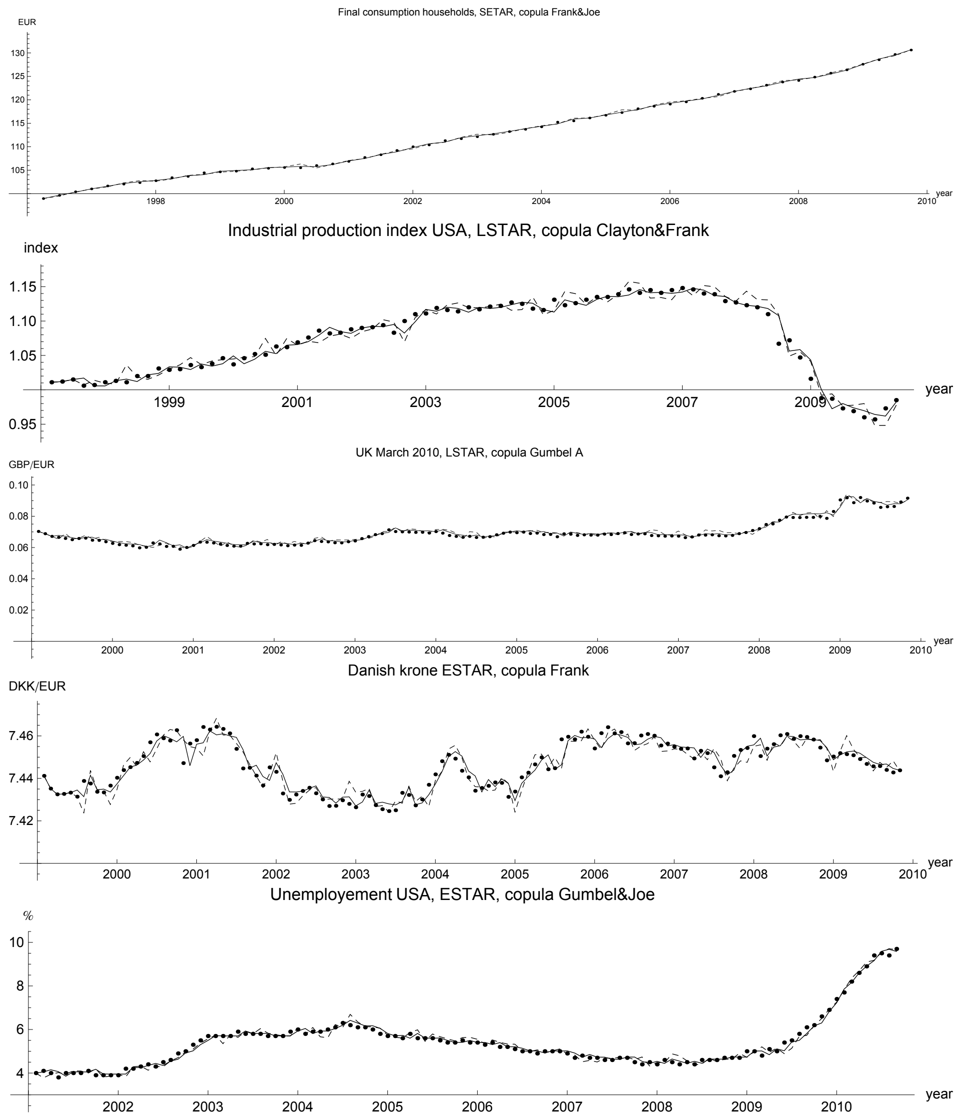

Fig. 1. Graphs comparing the original data (marked by dots) of 5 illustrative time series with their best model fits in both categories, the first (dashed line) and the second category (solid line).

of the research indicating that a linear model may forecast better than a nonlinear one, even when the nonlinear model is consistent with the actual data-generating process. For example, Montgomery et al. [16] show that a nonlinear model may forecast better than a linear one in some regimes, but not in others (e. g., recessions but not expansions). Similarly, Dacco and Satchell [5] show that a regime–switching model may have

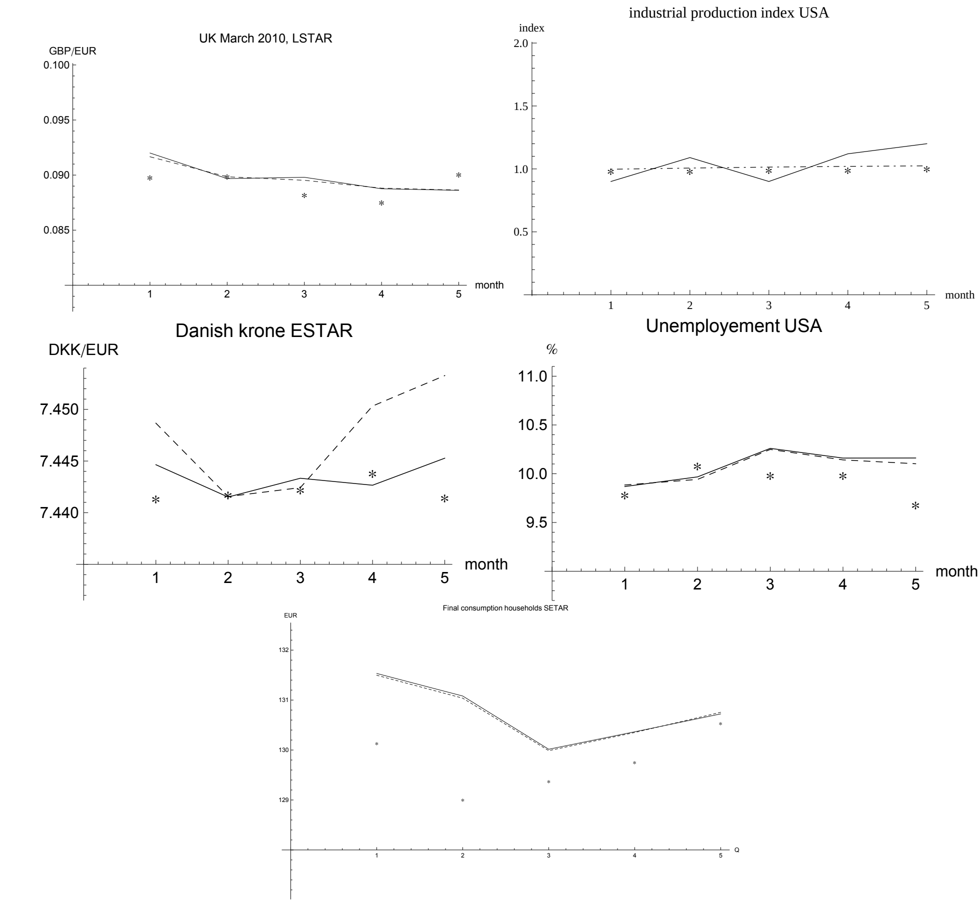

Fig. 2. Graphs comparing the predictions of 5 illustrative time series modeled by their best model fits in both categories with the last 5 original values.

poor forecasting performance relative to a linear model as a result of misclassifying observations. The point is that, any superior in-sample performance of a nonlinear model may not translate itself into superior out-of-sample performance. Moreover, forecasting with nonlinear models can be quite programming-intensive, as multistep-ahead forecasts need to be simulated."

From the optional models for 31 considered time series, the Table 7 summarizes the frequencies of occurrence of different candidates of Archimedean copulas, their convex combinations (denoted by &) and extreme value copulas.

4. CONCLUSIONS

Rakonczai [20] used autocopulas only for the testing of residual independence. Inspired by him we have extended his approach and applied autocopulas to the time series of the

Copula family	Number
Archimedean copulas (AC)	7
convex combinations of AC	13
extreme value copulas	11

Tab. 7. Numbers of the optimal copulas.

above mentioned residuals of the time series (exhibiting nonlinear dependences among subsequent residuals) in order to gauge how much they violate the assumptions of independence. We have tried to describe the nonlinear dependencies of the (time lag 1) residuals of the selected models by means of copulas (Archimedean, their convex combinations and EV copulas).

Then we have compared description of 'original' and 'modified' models. We have seen that when we have used copula driven noise series instead of the white noise in the models, the descriptions have been much better (in almost all cases). The average improvement in description has reached 51.42 %.

We have calculated also predictions by this copula approach but it comes to the small downgrade (in average −11.81 %). It is necessary to remark that most nonlinear techniques give good in-sample fits to real time series, but they are usually outperformed by models of random walks or random walks with drift when are used for forecasting [5].

This approach is new, so it is necessary to build up a theory about it. It will be the part of our further research. In the future we also want to describe real time series with non-Archimedean copulas like Gauss and Student copulas, Archimax copulas, etc. We also want to use regime-switching model with regimes determined by unobservable variables and compare it with the others.

ACKNOWLEDGEMENT

This work was supported by Slovak Research and Development Agency under contract No. LPP-0111-09, APVV-0073-10 and by VEGA 1/0143/11.

(Received July 29, 2011)

REFERENCES

[1] B. Akintug and P. F. Rasmussen: A Markov switching model for annual hydrologic time series. Water Resour. Res. *41* (2005).

[2] J. Arlt and M. Arltová: Finanční časové rady. Grada Publishing a. s., Praha 2003.

[3] W. A. Brock, W. D. Dechert, J. A. Scheinkman, and B. Le Baron: A test for independence based on the correlation dimension. Econometr. Rev. *15* (1996), 197–235.

[4] S. D. Campbell: Specification Testing and Semiparametric Estimation of Regime Switching Models: An Examination of the US Short Term Interest Rate. Department of Economics, Brown University, 2002.

[5] R. Dacco and S. Satchell: Why do regime-switching models forecast so badly? J. Forecast. *18* (1999), 1–16.

[6] P. Embrechts, F. Lindskog, and A. McNail: Modeling dependence with copulas and applications to risk management. In: Handbook of Heavy Tailed Distributions in Finance (S. Rachev, ed.), Elsevier *8* (2001), pp. 329–384.

[7] W. Enders and R. Pascalau: Pretesting the Out-of-sample Forecasting Properties of STAR Models. `http://www.faculty.plattsburgh.edu,2010`.

[8] P. H. Franses and D. Van Dijk: Non-linear Time Series Models in Empirical Finance. Cambridge University Press, Cambridge 2000.

[9] C. Genest and A. C. Favre: Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrolog. Engrg. (2007), 347–368.

[10] K. Ghoudi, A. Khoudraji, and L. P. Rivest: Propriétés statistisques des copules de valeurs extremes bidimensionnelles. Canad. J. Statist. *26* (1998), 1, 187–197.

[11] C. W. J. Granger and T. Teräsvirta: Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford 1993.

[12] S. Grønneberg and N. L. Hjort: The copula information criterion. In: Statistical Research Report *7*, Dept. of Math. University of Oslo, 2008.

[13] J. D. Hamilton: A new approach to the economic analysis of nonstationary time series subject to change in regime. Econometrica *57* (1989), 357–384.

[14] H. Joe: Models and Dependence Concepts. Chapman and Hall, London 1997.

[15] J. Komorník and M. Komorníková: Copula approach to residuals of Markov–Switching models for economic time series. In: Proc. AGOP 2011, pp. 39–43.

[16] A. Montgomery, V. Zarnowitz, R. Tsay, and G. Tiao: Forecasting the U.S. unemployment rate. J. Amer. Statist. Assoc. *93* (1998), 478–493.

[17] R. B. Nelsen: An introduction to copulas. In: Lecture Notes in Statistics *139*, Springer-Verlag, New York 1999.

[18] A. Prokhorov: A goodness-of-fit test for copulas. In: MPRA Paper *9998*, 2008, online at `http://mpra.ub.uni-muenchen.de/9998`.

[19] P. Rakonczai: On modeling and prediction of multivariate extremes with applications to enviromental data. In: Licentiate Theses in Mathematical Sciences (2009), pp. 83–109.

[20] P. Rakonczai, L. Márkus, and A. Zempléni: Autocopulas: Investigating the interdependence structure of stationary time series. Methodology Comput. Appl. Probab. *14* (2012), 1, 149–167.

[21] T. Teräsvirta: Specification, estimation, and evaluation of smooth transition autoregressive models. J. Amer. Statist. Assoc. *89* (1994), 208–218.

[22] T. Teräsvirta: Forecasting economic variables with nonlinear models. In: Handbook of Economic Forecasting (G. Elliot, C. Granger, and A. Timmermann, eds.), Elsevier *8*, Amsterdam 2006.

[23] H. White: Maximum likelihood estimation of misspecified models. Econometrica *50* (1982), 1–26.

Anna Petričková, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava. Slovak Republic.
e-mail: petrickova@math.sk

Magda Komorníková, Faculty of Civil Engineering, Slovak University of Technology, Radlinského 11, 813 68 Bratislava. Slovak Republic.
e-mail: Magdalena.Komornikova@stuba.sk